

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

## STAFF USE ONLY

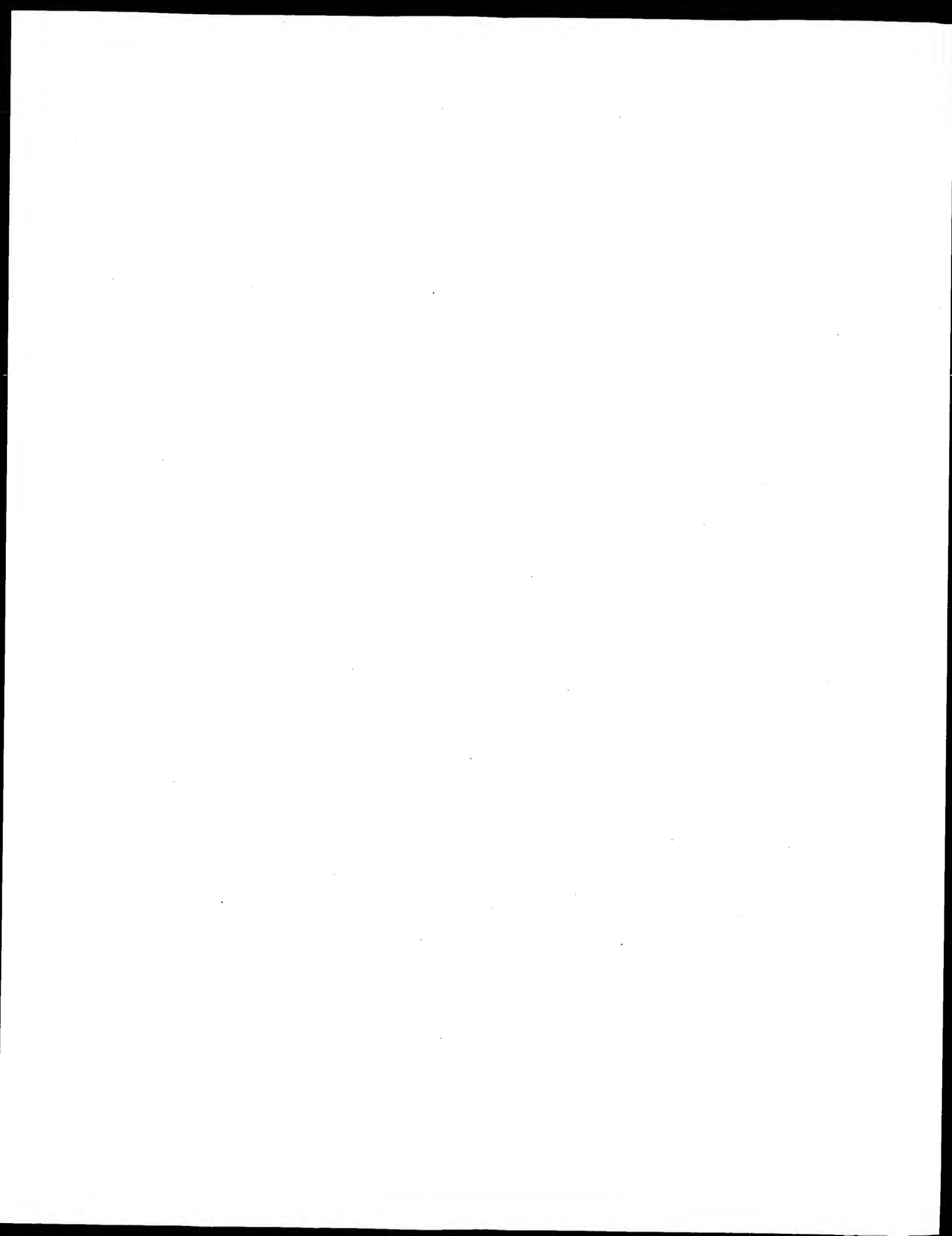
Searcher: P. Schreiber  
 Searcher Phone #: 308-4292  
 Searcher Location: CM112E18  
 Date Searcher Requested: \_\_\_\_\_  
 Date Completed: 11/14  
 Searcher Prep & Review Time: 6  
 Client Prep Time: \_\_\_\_\_  
 Filing Time: 10

## Type of Search

NA Sequence (#) 3  
 AA Sequence (#) 1  
 Structure (#) \_\_\_\_\_  
 Bibliographic \_\_\_\_\_  
 Litigation \_\_\_\_\_  
 Fulltext \_\_\_\_\_  
 Patent Family \_\_\_\_\_  
 Other \_\_\_\_\_

## Vendors and cost where applicable

STN \_\_\_\_\_  
 Dialog \_\_\_\_\_  
 Questel Orbit \_\_\_\_\_  
 CompuLink \_\_\_\_\_  
 Lexis Nexis \_\_\_\_\_  
 Sequence Systems CompuLink  
 WWW Internet \_\_\_\_\_  
 Other (specify): \_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 09:53:10 ; Search time 1462.83 Seconds  
(without alignments)  
12355.404 Million cell updates/sec

Title: us-09-589-510-3

Perfect score: 1912  
Sequence: 1 acccagcgctccgcaattt.....aaaaaaaaaaaaaaaa 1912

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_hic:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_inv4:*
207: em_gss_inv5:*
208: em_gss_inv6:*
209: em_gss_inv7:*
210: em_gss_inv8:*
211: em_gss_inv9:*
212: em_gss_inv10:*
213: em_gss_inv11:*
214: em_gss_inv12:*
215: em_gss_inv13:*
216: em_gss_inv14:*
217: em_gss_inv15:*
218: em_gss_inv16:*
219: em_gss_inv17:*
220: em_gss_inv18:*
221: em_gss_inv19:*
222: em_gss_inv20:*
223: em_gss_inv21:*
224: em_gss_inv22:*
225: em_gss_inv23:*
226: em_gss_inv24:*
227: em_gss_inv25:*
228: em_gss_inv26:*
229: em_gss_inv27:*
230: em_gss_inv28:*
231: em_gss_inv29:*
232: em_gss_inv30:*
233: em_gss_inv31:*
234: em_gss_inv32:*
235: em_gss_inv33:*
236: em_gss_inv34:*
237: em_gss_inv35:*
238: em_gss_inv36:*
239: em_gss_inv37:*
240: em_gss_inv38:*
241: em_gss_inv39:*
242: em_gss_inv40:*
243: em_gss_inv41:*
244: em_gss_inv42:*
245: em_gss_inv43:*
246: em_gss_inv44:*
247: em_gss_inv45:*
248: em_gss_inv46:*
249: em_gss_inv47:*
250: em_gss_inv48:*
251: em_gss_inv49:*
252: em_gss_inv50:*
253: em_gss_inv51:*
254: em_gss_inv52:*
255: em_gss_inv53:*
256: em_gss_inv54:*
257: em_gss_inv55:*
258: em_gss_inv56:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



[illegible]

DEFINITION	LOCUS	AK012170	RESULT 1
Mus musculus 10 days embryo CDNA, RIKEN full-length enriched library, clone:261028G06, full insert sequence.	1595 bp	mRNA	08-FEB-2001

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus  
1 (sites)  
REFERENCE  
Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning

Query Match	31.2%	Score 597,	DB 192,	Length 1595;
Best Local Similarity	66.5%;	Pred. No. 7e-138;		
Matches 916;	Conservative	0;	Mismatches 455;	Indels 6;
				Gaps 4.

	QY	DB	QY	DB
94	atgagagatcagagagaggtgcagtcagacttgaanaaagcagcagatctgcacccacacccac	153	154	atcaagaagcactcgccctcgaaagcaatagagatgtagcgttcgltgcgagcgaggtttctg
111	ATGAGAGATTGAGAGAGTGAATAGACCCACCAAAACGCGATCGCTCCACACGCAC	113	114	GTGAAGAGGCTTGGGGCTGGACGAAAGCGCGCTGAGCCAAACGAGCGGCTTCGGGTCGTG

OY	214	ggccagaggggagcgcgccgagggcgccgggcttgccggtgcacatgatcctgcgaagaag	27
Db	174	ggccagggaaagcggagagggcgatggcgctcatgttgatgaattaatcaaaacgacgaa	233
OY	274	atgagccggccggcggtgctcccttgagggtccggcccgccacggggaagaagcgctagc	333
Db	234	atggctggaaagactgctctgtttgggggggacctcttggaaactggccaaacacgcttgacc	293
OY	334	ctgggcttagcccaagagcttcggcagcaaggtcccttcgtccatggttgaagtcagaa	393
Db	294	cttccctatgtctcagggcgctgggacataaagctcccttctcccgatgggtgagcgag	353
OY	394	gtgtactccctcggaggttcaagaaacttgaggtgctgataggaaatttcggtagaagctata	453
Db	354	gtatactcaactgtagatcaaaaaagacagagtgctgattgagagaaatttcggagagccatt	413
OY	454	ggtttggtataaaggaacaaagaggtttataaggaagaggttactgaacttccca	513
Db	414	gggttggtggatgaaagagagctaaagaggttttatgaaggsgagtgacagagctcacctcc	473
OY	514	gaagaggtctgaagtagcaactggttgatatacgaagaagcattagccatgtaacatcagc	573
Db	474	tgtgagacagagaaaccccatgggtgggtracggcaaaatmncacccatggtgataaggg	533
OY	574	ttaagactgtttaaaggagcaactgaagtagttagatcttcataattgaatgctcgt	633
Db	534	ctcaagactgctaaaaagaaaccaaacagctgaagctggaccacccagatattttgaaacttg	593
OY	634	atcaagaaaaggctggcagttggtatgtttatatacagcaagcaaatgtagcagctg	693
Db	594	cagaaagaaagagtagagcgtggagattgtgatttaccttaaaccaaatgtatggagctgg	653
OY	694	aaaagagttgttagatgtagattcctttgtctacagaaacagatcttgaagctgaagtagat	753
Db	654	aaagagcaagggaggtgtgacacactatgcccacagactgtttgacctgaactgaagtagatc	713
OY	754	gttccatccccaagggtggaagttcatatagaaaaaagaaatctgtcgaagtagtcaact	813
Db	714	gtccct- ttggccaaagagatgtgcacaaag- aaagaaatattcatpaaagattgtgacctg	771
OY	814	catgaccttgatgcagcaaatgtctcagcccaaggtggcccaagatatcttgcctcttg	873
Db	772	catgacctgggatg- tgcataatgcgcggcctcagaggtgggcaagatatttgcctatgatg	830
OY	874	ggccagatgataagaaaccccgaaaactcgaataaccgaaaactacgccaagaataat	933
Db	831	ggccaggttgatgaagccaaaamaagcagaatcacagataaacttcgaaggagatcaaac	890
OY	934	aaggtggtanaatagatalatgcataaggaatgcagaactgttaacctgtgtttgtc	993
Db	891	aaagtggtgaacaaatcatattgacacagaggtgtggccgacctgtgtccctggagtgctttt	950
OY	994	attggtgggtgcacatggttgatatacgaatggttttcttactcaacgtgcatgag	1053
Db	951	gttcattgaggtccacatgctgatatccgattgcgtttacctactccacgagaccttgag	1010
OY	1054	agccatatalaccaatcgtgatacttgcaaatagaaggaatgtgaatgtagtagga	1113
Db	1011	tcttccatccgcccccatgttcatttttgatccacacagagcaactgtgtatagggggc	1070
OY	1114	act--- gatatgacaagtccaactggtataccggtgtgactctctagataggtcgtglat	1170
Db	1071	actgagacatcacatttctccacacggccatccctgttgatgtgctgaagagggatgatc	1130
OY	1171	attcgacacagacataatgccccctactgtagataagagatattgctatcccgagcaaa	1230
Db	1131	atcagagacacatgcatatnacggccacagagatgaagcagatcattaagatcagaccga	1190
OY	1231	gtgagagaggttatatgataagaaagacttctgttattttaggcggaatcgagacagcag	1290
Db	1191	acggaaagcatcaacatcagcagagagggcgtaaacacctctggggagattggccaccga	1250
OY	1291	acactcttgagacatgctatcaatttgatatacactgcccagctgtgtcctaaagactaat	1350

Db	1251	ACCACCTCTGAGGTATTCGTGTGACAGCTCTGTACCCCACTCCTGTGGCAAAAGTCAAT	1310
Qy	1351	ggaagaagaagaatctgcaagcgtgatctcgaagaaatcagtggcctcattgatatgc	1410
Db	1311	GGCAGGACACACATTGAGAAAGGAGCACCGTGAAGAAATCAGGACCTCTTCTATGACGCC	1370
Qy	1411	aaatctcgcgtcgcgcgtccagagcaagaagaataatactaactgaatttg	1467
Db	1371	AAGTCCTCTGCAAGATTTTGGCTGTACCAAGACAGACACAGTACATCAAGTAACGTTGG	1427

RESULT	2
BE512278	
LOCUS	BE512278 610 bp mRNA EST 07-AUG-2000
DEFINITION	946086D09.y1 946 - fassel primordialium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION	BE512278
VERSION	BE512278.1 GI:9733421
KEYWORDS	EST.
SOURCE	Zea mays.
ORGANISM	Zea mays.

REFERENCE

1 (bases 1 to 610)

male birds from various cDNA libraries sequenced at Stanford University  
Unpublished (1999)  
Contact: Walbot V

Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946066 row: D column: 09.

```

SOURCE
1. .610
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassell primordialium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/notes="Organ: tassels; Vector: HybriZAP; Site:1: EcoRI;
Site:2: xhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350
to 3 Kb with a 1 Kb average."
160 a 135 c 196 g 119 t

```

Query Match	30.9%	Score 591.6	DB 136	Length 610
Best Local Similarity	99.2%	Pred. No. 1,2e-13e		
Matches 605, Conservative	0	Mismatches 4	Indels 1	Gaps
2P	86	tggcgcgcatgtaggattcagagagtgacgtgacctcgaagaagcagcgtccaccc	145	
2b	2	TGGCGGCGAATGAGGATTCGAGAGATGTCAG-CGACCTCGAAGACAGCGCATCGCCACCC	60	
2Y	146	acaccaccatcaaggagactcgctctgaagcgcaatggatgtgcatgtgctltgcgcgcg	205	
2b	61	ACACCCACATCAAGGGAGACTGGGCTTCGAGCCAAAGGATGGCATTTGCTTGGCGGCGG	120	
2Y	206	gattcgtlgggcacgagcgcgcgcgcgagcgcgccgggctlggcggttcgacatgattcgcc	265	
2b	121	GGTTCTGGGGCCAGTGGCGGCGCGCGAGGCGCGCGGAGCTGGCGTGCACATGATTGGCT	180	

[illegible]



BASE COUNT	ORIGIN
197	a
127	c
182	g
210	t
	1 others

Query Match	25.9%	Score 494.6	DB 154	Length 717
Best Local Similarity	84.2%	Pred. No. 1.8e-112		
Matches 630	Conservative	1	Mismatches 76	Indels 41
				Gaps 5
QY 955	gatagaaggaatgcagagcttgtaacctgctgctgctgtgtttgttcaatgtaataaggtccacatgctg	1014		
Db 11	GAGGAGAGGAATTCGACGACCTTGTAACCTGGTGCTTTGTTCAATTGATGAGGCCACATGTTG	70		
QY 1015	gatacgaatgttttcttacttaacocggtcatcttggaagccatataccaatccgtg	1074		
Db 71	GATATCGAATGCTTTTCTTATCTTAACTGTCACCTGGAGACCCCTTATACCAATTTGG	130		
QY 1075	atactgtcacacaataagggaaatgtaaatgtaagaggaactgatalgacaagttccaat	1134		
Db 131	ATATCTGTCACCAATATAGGGAAATGATGTAATGTAAGAGAACTGATATGACAACTCCACAT	190		
QY 1135	ggtataccgggtgagctctctatagatagctgctgctgtgtttttcggaagaagacatctgccc	1194		
Db 191	GGTATTCACGATGATCTTCTTAATATGTTGGTGGATTTTCGACAGAAACATATAGGCGCT	250		
QY 1195	actgaagatatacagatalatgtgctatccgagacacaagctggagagagatgataatgata	1254		
Db 251	GCTGAGATGATACAGATATTGGCAATCCGACACAAAGGAGAGAGATGATATGATGAA	310		
QY 1255	gaaagctctgtctatttaagcgagatcggaaagcagacatcttligaacatgattca	1314		
Db 311	GAAGGCTTGCTTTATTTTAAAGCGAGATGGACAGACATCTTTGACATGCTATTCAG	370		
QY 1315	ttgatctacactgcagcgtgctcctcaagataatggaagagagaaatctgaagct	1374		
Db 371	TTGCTATACCTCTCCAGCGCTGGTGCAAAGACTAATGGAAGAAAGATCTGCAGAGCT	430		
QY 1375	gatacgaagaaagtcagctgggtctctatcttgatgccaatctctggctcggctctcag	1434		
Db 431	GATCTTGAAGAAATGTAGTGCTCTCTTTTGGATGCCAAATCTTCGGCTGTGCTCCAG	490		
QY 1435	gagaaacaagaagataacatacacttgatattgatctcctgctcggaagctcgaag	1494		
Db 491	GAGCAACAAAGCAATATACATCACTAGATT----TCTCTGTTGTAAGAAAGCTCGAAGAA	546		
QY 1495	aattagttgcagcgtcgaagaatcatctagtcattgatctcgtcttcaagtgctagtc	1554		
Db 547	AATGAGTGGCAGACTC---AAGGCCATGTAGATGATCTCTTCACATCTGGGGGAAAG	603		
QY 1555	tactggtcttggaagcagcagcatcttcggggggaagcgcttgaaatttgagtgacct	1614		
Db 604	TTATTTTGTATGTCACGCTTAT-----GGCTT	631		
QY 1615	gtgttagctccaragaagaacttgcttcggcatatgtcgttcaagcactgctgcgtg	1674		
Db 632	CGGTAGCTCCAGAGAGAACTTGGTCCGGTATATGCTTTCACACATA-TGTTGCTTN	690		
QY 1675	attagattgtccacgcggtggcaggaatg	1702		
Db 691	A-TAGATTGACAATTGGTCAGSHAATG	717		

RESULT 6

AM927557

AM927557

LOCUS

DEFINITION

512 bp

mrna

945018605.X2 945

Mixed adult tissues from Walbot lab, same as 707

EST 30-MAY-2000

ACCESSION	(SK) zea mays cDNA, mRNA sequence.
AM927557	
AM927557.1	GI:8102879
VERSION	EST.
KEYWORDS	zea mays.
SOURCE	zea mays
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; zea. 1 (bases 1 to 512)
REFERENCE	
AUTHORS	Walbot,V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL	Unpublished (1999)
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu plate: 945018 row: G column: 05. positional/clonalifiers

[illegible]

QY 650 cagtggtgtagtatacatcgaagcaaatagtcgaagtcgaagagtgtagat 709  
 DB 421 CAGTGGGTGATGTTATATACATGGAAGCAAAATAGTGACAGCTGAAAAGAGTTGATAGT 480  
 QY 710 gtatctctttgtacagaatacagctctgaa 741  
 DB 481 GTGATCTTTTGTCTACGAAATATGATCTTGAA 512

RESULT 7  
 BE512425 549 bp mRNA EST 07-AUG-2000  
 LOCUS 946071A06.x1 946 - tassel primordium prepared by schmidt lab Zea  
 DEFINITION mays cDNA, mRNA sequence.  
 ACCESSION BE512425  
 VERSION BE512425.1 GI:9733673  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 Clade; Panicoidae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 549)  
 AUTHORS Walbot, V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
 JOURNAL University  
 COMMENT Unpublished (1999)  
 CONTACT Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 946071 row: A column: 06.  
 FEATURES  
 source Location/Qualifiers  
 1..549  
 /organism="Zea mays"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /clone\_lib="946 - tassel primordium prepared by schmidt  
 lab"  
 /issue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XLOLR"  
 /note="Organ: tassels; Vector: HybridZAP; Site: 1: EcoRI;  
 Site: 2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybridZAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 kb average."  
 BASE COUNT 150 a 146 c 116 g 137 t  
 ORIGIN

Query Match 24.5%; Score 469; DB 136; Length 549;  
 Best Local Similarity 94.5%; Pred. No. 4.1e-106;  
 Matches 534; Conservative 1; Mismatches 11; Indels 19; Gaps 4;

QY 1210 atattggcattccgagcaagtcgaagtcgatatgataagaagtcgtctat 1269  
 DB 549 ATATTGGCATTCGAGCCAAAGTGAGGACATTTATGATGTAAGAAAGCTTTGCTTAT 490  
 QY 1270 ttaaggcagatcgacacagacatcttgagacatctatcaatgatatacctgc 1329  
 DB 489 TTAGCGGAGATCGACACACACATCTTTAAGATCTTATCAATTGATATCACTGCC 430  
 QY 1330 agcctggcctaaagactcaatgtagagagaatctgcaagctgatactcgaggaagtc 1389  
 DB 429 AGCGTGGCTTCAAGACTAATATGAGAGAAATATATCAAGCGCTGATCTCGAGGAAGTC 370  
 QY 1390 agtggcctatcttgatgccaatcctcgctcgctgctccagagacaagaaga 1449

DB 369 AGTGGCTCATTTTGGATGCCAAATCTCGGCTCGGCTGCCAGACCAAGAAAAGA 310  
 QY 1450 tacatacctagatttgatctctcgtgctgtagaagtcctgaagaatgtagtgcagc 1509  
 DB 309 TACATACCTTAATTTTGGATCTGCTGCTGAGAACTCGAAGAGAAATGATGTCAGC 250  
 QY 1510 tcgaagtcatactagtcgatactgatactgatacagtcacagtcatactgcttgagc 1569  
 DB 249 TCGAAAGTCATCTAGTGAATGATGATCTGCTACA-----GCTCTTGAGC 205  
 QY 1570 gacacatttcggggggggaagcgttgaaatttcagtcagtcgtctgtagtccacac 1629  
 DB 204 GAGCAATTTTGGGGGGGAGCGCTTGAAATTTTGACAGTCCGCTTGTTAGTCTCCAA 145  
 QY 1630 gaagacttgatccgacatctgctgctacagacgttcgctgattagattgctcacc 1689  
 DB 144 -AAACTTGTACCGCGCATATGCTGCTTCACGACCTGTCGCTGATGATGATGTCACC 86  
 QY 1690 gttgaggaattccggtgtgtg-tttatcttgatcaltcgtgctcgagtc--tgcct 1746  
 DB 85 GGTGAGGAATTTGCCGTGTGTTTATCTTCATCGTGTGCTCGGATCTGTGCT 26  
 QY 1747 ccacgggtgattggtgcccgaacc 1771  
 DB 25 CCACGGGTGATTTGCCCGAACCC 1

RESULT 8  
 BE512426 474 bp mRNA EST 07-AUG-2000  
 LOCUS 946071A06.y1 946 - tassel primordium prepared by schmidt lab Zea  
 DEFINITION mays cDNA, mRNA sequence.  
 ACCESSION BE512426  
 VERSION BE512426.1 GI:9733674  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 Clade; Panicoidae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 474)  
 AUTHORS Walbot, V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
 JOURNAL University  
 COMMENT Unpublished (1999)  
 CONTACT Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 946071 row: A column: 06.  
 FEATURES  
 source Location/Qualifiers  
 1..474  
 /organism="Zea mays"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /clone\_lib="946 - tassel primordium prepared by schmidt  
 lab"  
 /issue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XLOLR"  
 /note="Organ: tassels; Vector: HybridZAP; Site: 1: EcoRI;  
 Site: 2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybridZAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 kb average."  
 BASE COUNT 110 a 121 c 162 g 81 t  
 ORIGIN





12

\_\_\_\_\_

1





```

BASE COUNT      http://fulllength.invitrogen.com"
ORIGIN          274 a      188 c      279 g      175 t      3 others

```

	Query Match	Best Local Similarity	23.4%;	Score 448;	DB 106;	Length 919;	
	Matches 629;	Conservative	69.9%;	Pred. No. 7.9e-101;	2;	Mismatches 267;	Indels 2;
						Gaps	2;
OY	94	atgaagatcgaagaaggctgacgtccacctgaaagaacagcgatgatgccaccacaacccaac	153				
Db	20	ATGAAGATTGAGSAGGTGAAGAAGACACTACGAAGACCAAGCCATCGCTCCACAGCCAC	79				
OY	154	atcaagggactcg -gccttcagcgcaatlygaatggcatttgcttggcgcgagttcgt	212				
Db	80	GTTAAAGGGCTTGAGGCTTGACGACAGCGCGCTTGGCCAAG;AGGGGGCCTCGAGGGCTTGT	139				
OY	213	gagccagcgagcgcgcgcgcgagcgcgcgcgcggtgctgcgcatgatctgcagaaga	272				
Db	140	GGGCCAGGAACGCGCCAGAGGACATGTGGCTCATTAATAAATTAACAAGCAMAGA	199				
OY	273	galtagccgagcgcgsgltgctcttgcggtlccgcccagcgagcgagcgctlagc	332				
Db	200	AATGGCTGGAMAGACTGTCCTTGTGTGGCAGGACCTCTTGAACTGGCAAGACAGCTCTGGC	259				
OY	333	gcttcgcataagccccagcgagcttcgagcaagaagtcccttctgctoctatlyglatgatacga	392				
Db	260	TCTGGCTATTGCTCAGGAGCTGGGTAGTMAAGTCCCCCTTTCGCCCAATGGTGGGAGTGA	319				
OY	393	agtlgactccttcgagagtcgaagaanaactgaggtgctctaigtgaaaatttcgtagagctat	452				
Db	320	ACTTACTCACTGARATCAACAAAGAACAGACAGAGGTGCTGATGGAGAACTTCCGCAC -GCCAT	378				
OY	453	aggttgcgclataaagaaaaaagaaggtllatalygaagagaggttactgaacttccc	512				
Db	379	TGGGCTGGCAATTAAGGAGACCAAGGAAGTTTATGAAGGTGAAGTCAACAGAGCTAACTCC	438				
OY	513	agaagagcttaagagtiacaactcgttgatattgcaaaaagaattagcacatgtaatcag	572				

Db	439	GTGTGAGACAGAGAAATCCCATGGGAGGATATGGCAAAACCAATTAGCCATGTGATCATAGG	498
Qy	573	cttaaaagactgtlaaaggagactaaagcaactgaagtlagatcttccaattatgaatgcct	632
Db	499	ACTCAAAACAGGCCAAGGACCAACCAACGATTGTAACCTGGACCCGACGATTTTGGAAAGTTT	558
Qy	633	gattcaaggaagaagtgaggaatgggttgatgttatataatcgaagcaaatagtggagcgt	692
Db	559	GCAGAAAGCGCAGTAGAGAACTGGAGATGTGATTTACATTTGAAGCCACACAGTGGGGCGGT	618
Qy	693	gaaaagatltgtagatgtgaattccttcgtccagaatacgaatccttgaagctgaagaata	752
Db	619	GAAGAGGACGAGGACAGGTGTGATACCTATGGCACAGATTTCGACTTTGAAGCTGAAGAGATA	678
Qy	753	tgcttcctatccccaaggtgaagtlcaatlaaagaaaaagaatttgcagatgylcacact	812
Db	679	TCCTTCCCTTTGGCAAAAGGGGATGTGCACAAAAGAAAGAAATCATTCACAAATGTACCTT	738
Qy	813	tcatgaacttgatgcagcaaatgtcgaagacagaggtgaccagaatatattgtccctat	872
Db	739	GCATCATCTTGGATGTGGCTTATGTGGGGGCCCCAGGGGGACAAAGATATCTGTGCATGAT	798
Qy	873	gggccagatltgaacacacgaagaaactgaatcgcgcgaagaaacacgaagcaagaattaa	932
Db	799	GGGCCAGCTTATGAAAGCCCAAGAAAGACAGAAATCAGACAAACTTCAGAGGGGAGATTAA	858
Qy	933	taagtgltgataatgatataatcgaatgaagaagatcgcgaagcttgcacctgtgttcctgt	992
Db	859	TAAAGTGGTGAACAAAGTACATCCAGCAGGACATTTCTGACGTGAGTCCGGGTGTCTGTT	918

LOCUS	AL557272	962 bp	mRNA	EST	16-FEB-2001
DEFINITION	AL557272 INT. FL012_T01	Homo sapiens	cDNA clone	CSDBH004YM17	5 prime
ACCESSION	AL557272				
VERSION	AL557272.1	GI:12900714			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 962)				
JOURNAL	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization				
	Unpublished (2001)				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.				
FEATURES	Location/Qualifiers				
source	1..962				

Query Match	22.7%	Score 433.2;	DB 106;	Length 962;
<p>/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life technologies. Contact : Feng Liang Life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (41) 301 610 8371 Email : <a href="mailto:filang@lifetech.com">filang@lifetech.com</a> URL : <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a>"</p>	278 a	201 c	295 g	185 t 3 others



Wed Nov 14 08:34:11 2001

us-09-589-510-3.std.rst

Page 14

```

OY 1571 agacacatttcggggg-----aaagcgcttgaatttgcagtgccgc-ctgtgtttagt 1622
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 AGCACGTTTCGCGCGGSAAGCAAAACTGTTGAAATTTGTATGACCTGTTTGTGAGC 10
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1623 ctctcarraga 1631
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 CTCACGCGCA 1
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: November 13, 2001, 11:07:51  
Job time: 4481 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2001, 06:55:54 ; Search time 18.34 Seconds  
(without alignments)  
558.288 Million cell updates/sec

Title: US-09-589-510-4  
Perfect score: 2263  
Sequence: 1 MREEVOSTSKKRIATHTH.....YLDKSSARLLQEQERYIT 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1721.5	76.1	456	2	US-08-933-750C-5
2	1721.5	76.1	456	4	US-09-234-613-5
3	117	5.2	806	4	US-08-999-774A-6
4	111.5	4.9	327	4	US-09-446-504-3
5	111.5	4.9	652	4	US-08-987-123-2
6	108.5	4.8	686	2	US-08-993-228-12
7	104.5	4.6	273	4	US-09-928-213B-11
8	104.5	4.6	942	3	US-09-074-579-1
9	104.5	4.6	942	4	US-09-388-774-1
10	101.5	4.5	273	4	US-08-928-213B-10
11	100.5	4.4	1184	4	US-09-541-782-2
12	98	4.3	2101	1	US-08-466-390-4
13	98	4.3	2101	1	US-08-470-950-4
14	98	4.3	2101	1	US-08-467-781-4
15	98	4.3	2101	2	US-08-483-924-4
16	98	4.3	2101	4	US-09-452-294-1
17	97.5	4.3	456	4	US-09-268-364-21
18	97	4.3	2101	1	US-08-195-487-4
19	97	4.3	2101	5	PCT-US93-06160-4
20	96.5	4.3	876	1	US-08-717-515-4
21	96.5	4.3	1276	1	US-08-717-515-8
22	96	4.2	544	3	US-08-559-397A-29
23	95.5	4.2	1018	1	US-08-072-610-2
24	95.5	4.2	1018	2	US-08-719-822B-2
25	95.5	4.2	1018	4	US-09-092-458-2
26	94.5	4.2	1119	4	US-09-396-651B-2
27	93.5	4.1	479	4	US-09-446-504-64

28	93.5	4.1	545	2	US-08-467-822-30	Sequence 30, Appl
29	93.5	4.1	545	4	US-08-432-697-30	Sequence 30, Appl
30	93.5	4.1	545	4	US-08-466-248-30	Sequence 30, Appl
31	93	4.1	405	2	US-08-222-719-2	Sequence 2, Appl
32	93	4.1	405	2	US-08-470-925-2	Sequence 2, Appl
33	93	4.1	405	2	US-08-471-613-2	Sequence 2, Appl
34	93	4.1	405	5	PCT-US93-10443-2	Sequence 1, Appl
35	92.5	4.1	406	2	US-08-222-719-1	Sequence 1, Appl
36	92.5	4.1	406	2	US-08-470-925-1	Sequence 1, Appl
37	92.5	4.1	406	2	US-08-471-613-1	Sequence 1, Appl
38	92.5	4.1	406	5	PCT-US93-10443-1	Sequence 3, Appl
39	92.5	4.1	1964	2	US-08-790-912-3	Sequence 2, Appl
40	92.5	4.1	2052	2	US-08-790-912-2	Sequence 2, Appl
41	90.5	4.0	249	2	US-08-991-946A-1	Sequence 1, Appl
42	90.5	4.0	571	4	US-08-961-083-4	Sequence 4, Appl
43	90.5	4.0	834	1	US-08-471-033-21	Sequence 21, Appl
44	90.5	4.0	834	2	US-08-471-044-21	Sequence 21, Appl
45	90.5	4.0	834	2	US-08-463-483A-21	Sequence 21, Appl

## ALIGNMENTS

RESULT 1  
US-08-933-750C-5  
Sequence 5, Application US/08933750C  
Patent No. 5932442  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guebler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HMCINOT01

CLONE: 9476  
US-08-933-750C-5

Query Match 76.1%; Score 1721.5; DB 2; Length 456;  
Best Local Similarity 72.5%; Pred. No. 1.3e-154;  
Matches 330; Conservative 72; Mismatches 52; Indels 1; Gaps 1;

```

OY 1 MKEIEVOSTSKKRIATHTHIKGLDANGMAIALAGFVGGAAREAGLAVDMIRKK 60
DB 1 MKEIEVSTTKTORIASHSHVKGGLDESLAKQAASGLVGENARACGYIELLESKR 60
OY 61 MAGRAVLLAGPAPGKTALALAGIAGELSKVPCPCPMGSEVYSSEVKT EYLMENFRRAI 120
DB 61 MAGRAVLLAGPAPGKTALALAGIAGELSKVPCPCPMGSEVYSSEVKT EYLMENFRRAI 120
OY 121 GLRIKENKEVEGEVTELSPEEASTTGGAKSISHVILSKTKVKKOKLIDSSYDAL 180
DB 121 GLRIKETKEVEGEVTELTPECTENPMGCGIKTISHVILGLTKAKGKOLKIDPSIFESL 180
OY 181 IREKAVADVITYIANSAGAVKRGCDSPATEYDLEAEYVPIPKGEVHKKEIVODVTL 240
DB 181 QKEVEADGVITYIANSAGAVKRGCDTYATEFDLEAEYVPLPKGDVHKKEIIVODVTL 240
OY 241 HDLDAANAPOGGODILSLMGOMKPKTEITTEKLOEIKKVVNRIYDEGIAELVPGVLF 300
DB 241 HDLDAANARPOGGODILSMGOLMKPKTEITDKLGEIKKVVNKYIDOGIAELVPGVLF 300
OY 301 IDEVHMLDIECFSTYLRALSPSPVILATNRGICNVRGT-DMTSPHGIPVLLDRVI 359
DB 301 VDEVHMLDIECFSTYLRALSPSPVILATNRGICNVRGTEITSPHGIPVLLDRVMI 360
OY 360 IRTETGPTETIOILAIRAOVEEIDMDESLAYIGEIGQTSLRHAIQLISPAVSYSKTN 419
DB 361 IRTMLTYQEMKQIKIRAOEGINISEALNHLGEIGTKTLRSYVQLTTPANLLAKIN 420
OY 420 GREKICKADLEEVSGLYLDAKSSARLLQEOOERYI 454
DB 421 GKDSIEKEHVEISLSELFYDAKSSAKILLADOODKYM 455

```

RESULT 2  
US-09-234-613-5  
Sequence 5, Application US/09234613  
Patent No. 6132973  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HMCINOT01  
CLONE: 9476  
US-09-234-613-5

Query Match 76.1%; Score 1721.5; DB 4; Length 456;  
Best Local Similarity 72.5%; Pred. No. 1.3e-154;  
Matches 330; Conservative 72; Mismatches 52; Indels 1; Gaps 1;

```

OY 1 MKEIEVOSTSKKRIATHTHIKGLDANGMAIALAGFVGGAAREAGLAVDMIRKK 60
DB 1 MKEIEVSTTKTORIASHSHVKGGLDESLAKQAASGLVGENARACGYIELLESKR 60
OY 61 MAGRAVLLAGPAPGKTALALAGIAGELSKVPCPCPMGSEVYSSEVKT EYLMENFRRAI 120
DB 61 MAGRAVLLAGPAPGKTALALAGIAGELSKVPCPCPMGSEVYSSEVKT EYLMENFRRAI 120
OY 121 GLRIKENKEVEGEVTELSPEEASTTGGAKSISHVILSKTKVKKOKLIDSSYDAL 180
DB 121 GLRIKETKEVEGEVTELTPECTENPMGCGIKTISHVILGLTKAKGKOLKIDPSIFESL 180
OY 181 IREKAVADVITYIANSAGAVKRGCDSPATEYDLEAEYVPIPKGEVHKKEIVODVTL 240
DB 181 QKEVEADGVITYIANSAGAVKRGCDTYATEFDLEAEYVPLPKGDVHKKEIIVODVTL 240
OY 241 HDLDAANAPOGGODILSLMGOMKPKTEITTEKLOEIKKVVNRIYDEGIAELVPGVLF 300
DB 241 HDLDAANARPOGGODILSMGOLMKPKTEITDKLGEIKKVVNKYIDOGIAELVPGVLF 300
OY 301 IDEVHMLDIECFSTYLRALSPSPVILATNRGICNVRGT-DMTSPHGIPVLLDRVI 359
DB 301 VDEVHMLDIECFSTYLRALSPSPVILATNRGICNVRGTEITSPHGIPVLLDRVMI 360
OY 360 IRTETGPTETIOILAIRAOVEEIDMDESLAYIGEIGQTSLRHAIQLISPAVSYSKTN 419
DB 361 IRTMLTYQEMKQIKIRAOEGINISEALNHLGEIGTKTLRSYVQLTTPANLLAKIN 420
OY 420 GREKICKADLEEVSGLYLDAKSSARLLQEOOERYI 454
DB 421 GKDSIEKEHVEISLSELFYDAKSSAKILLADOODKYM 455

```

RESULT 3  
US-08-999-774A-6  
Sequence 6, Application US/08999774A  
Patent No. 6274312  
GENERAL INFORMATION:  
APPLICANT: Gish, Kurt C.  
APPLICANT: Seghezzi, Wolfgang  
APPLICANT: Shanahan, Frances  
APPLICANT: Lees, Emma M.  
TITLE OF INVENTION: Intracellular Regulatory Molecules;  
TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:

Page 3

```

      PRIOR APPLICATION NUMBER: JP 9-187496
      PRIOR FILING DATE: 1997-06-26
      PRIOR APPLICATION NUMBER: JP 9-320692
      PRIOR FILING DATE: 1997-11-27
      NUMBER OF SEQ ID NOS: 92
      SOFTWARE: PatentIn Ver. 2.1
      SEQ ID NO 3
      LENGTH: 327
      TYPE: PRT
      ORGANISM: Pyrococcus furiosus
      US-09-446-504-3

Query Match          4.9%; Score 111.5; DB 4; Length 327;
Best Local Similarity 21.5%; Pred. No. 0.017;
Matches 79; Conservative 63; Mismatches 112; Indels 113; Gaps 18.

QY      54 DMIRKKMAGRA-----VLLAPRATGKALALGIAOELSGSKVPFCPCWGS--- 99
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       26 DIVGOEIVIRLKHVYKGTGSMPLLLPAGPGVGGKTTAALLAREL-----FGENMR 76
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      100 -----EYVSSEVKKTEVLM-----NFRRAIGLRITKENKEYVEGEVTELSPEAE-----TTGG 149
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       77 HNFELNASBERGINIAREVKKEFARTKPIGGSFKIIFLDEADALTDQAQALRTTEM 136
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      150 YAKSISHVILSLTVKGTOKLIDISY-----DALIKKXAVAGDVIYIANGSAVKRV 203
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       137 FSNVFRFLSCNYSKKIIEPIOSRCATFFRRPRLRDEDIARKLR-----YIANEGLE--- 188
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      204 GRCSFATPEVDLEAEVVPPIPKGEVAKKKEIYODVTLHD-----LDAANQPOGGQ 254
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       189 -----LTRECLQILY--IAEGDMRRAINILLOAAALDKRTDENVPVMSRARPEDIR 240
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      -----LRECLQILY--IAEGDMRRAINILLOAAALDKRTDENVPVMSRARPEDIR 312

```

[illegible]

```

Db 186 FSCVAGAEKEKELVEVEFLKDPKRTKLGARIPAGVILGPPCTGKTLAKAVAGAG 245
OY 89 SKVPCPMVSE-----VSESVKTEVLMENFRA-----IGRIKENVEVEGEVTE 137
Db 246 --VPEFSIGSPFVEKFEVVGASRVR--LEEDAKKAPAIIFDLNDVAGRGORGGLG 300
OY 138 LSPFEAEESTTGVAKSISHVITSLKTVGTRQKLKSSIDALIKKAVAGDVITYIBANS 197
Db 301 GGNDEREQT-----LNOLLLEMDGFEQNE-----GIYIATNR 334
OY 198 G-----AVRVGRGDS--FATEYDEAEVYPIRKEVH--KKETVODVTLHDDAAMAQ 249
Db 335 SVYLDPALRPEREDRKVLVGRPDVKGREAIL--KVHAKNKPLADV--DLKLVAQ 387
OY 250 POG--GODILSLMGQ--MMKPKTEITEKLRQEIKNV--NRYIDEGIAELVGVLFIDE 303
Db 388 TGFVGADLEVLNMAALVAAR-----NKSIIIDASIDIDEADVIAGPSKKDK 436
OY 304 -VHMDICFSTLNRALESPLVILATNRGICNV--RGTDMTSPHGIP----- 350
Db 437 TVSOKERELVAV--HEAGHTIVG--LVLSARVVKTVIVPRGAGVMIALPREDMLLS 493
OY 351 -VDLDR-----VIETETYG-----PREMOIILAIRQVEIDMDESIA 391
Db 494 KEDMKQGLAGLMGGRVAEELIFNVOTTGASNDPEQATOM--ARMVTEYMSER-- 545
OY 392 YLGEIGQOOTSIRHA--IOLISPAVSASTNGREKICKADLEVSGLYLAKS--SARLLOQ 449
Db 546 -LGVP--OYEGNHAMLAGSPKQISBQTAVE--ID--EVVSLNENRKNAAELIQSN 597
OY 450 QERY 453
Db 598 RETH 601

```

RESULT 6  
 US-08-993-228-12  
 ; Sequence 12, Application US/08993228  
 ; Patent No. 5976838  
 GENERAL INFORMATION:  
 APPLICANT: Jacobs, Kenneth  
 APPLICANT: McCoy, John M.  
 APPLICANT: Lavallee, Edward R.  
 APPLICANT: Racie, Lisa A.  
 APPLICANT: Merberg, David  
 APPLICANT: Treacy, Maurice  
 APPLICANT: Spaulding, Vikki  
 APPLICANT: Agostino, Michael J.  
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genetics Institute, Inc.  
 STREET: 87 Cambridgepark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/993.228  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sprunger, Suzanne A.  
 REGISTRATION NUMBER: 41,323  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 498-8284

```

; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-228-12

```

```

Query Match 4.8%; Score 108.5; DB 2; Length 686;
Best Local Similarity 20.1%; Pred. No. 0.1;
Matches 109; Conservative 74; Mismatches 184; Indels 175; Gaps 25;

```

```

OY 23 GLGLDNGMALAALAAFGQAAA-----REAGLA-----VDMIRCK-- 60
Db 185 GIGRTGRMGGLFS--VGELTAKVLKDEIDVFKDVAGCEAKLEIMEVFNKPKOY 241
OY 61 -----MAGRAVLVAGPATKKTALALGIAOELGSKVPECPMGSEVYSEVKTEVLMEN 115
Db 242 QDLGAILPKGAILTGPPGTGKTLAKATAGE--ANVPFITVSGSE-----FLEM 288
OY 116 FRAAIGLRKE-----NKE-----VYEGEYVELSPEAEESTGTYAK----- 152
Db 289 FVGGRPARVRDLFLAKKNAPCLFIDEIDAVGRKKGRNGFGSQSEPTLNOLLVEMDG 348
OY 153 --SISHYIISLKT-----VKG-----TKQLKDSIYDALIKKAVAV-----G 188
Db 349 FNTTNNVILAGTNRPGPPDIKGRASIFKVHLRPLKLDSTLEKDKLARKLASLTFGFGSA 408
OY 189 DVITYEANSQAVKRVGRCDSFATEYDLAEVYPIRKEVHKKELVQ-----DVLHDL 243
Db 409 DVANVCNEALILAAHLSDSINQKHFEQAIERY--IGGLKRTYLOPEEKRTVAHYHA 465
OY 244 DAANAQ-----POG-----GODILSLMGQMKPKRTE- 270
Db 466 GHVAVAGYLEHADPLLKVSIIIPRGKLGVAQYLPRKQOYLITREQLDKMCKMTIGRVSEE 525
OY 271 -----ITEKROEINKVNV--RYIDEGIAELVGVLFIDEVHMDICTSYNRALES 321
Db 526 IFGRITTGADDLKRVTSAYAAQIVQFGMNEKV--GOISFDLPBGDM-----VLEK 576
OY 322 PLSPIVLATNRGICNVRGTDMTSPHGIPVDLDR-----LVITRTETGPTMIO 372
Db 577 PYSE-----AFARLIDDEVAILINDAYKRTVALLTEKKADVEKVALLEKEVLKNDIVE 632
OY 373 ILAIRQVEIDMDESILAYLGEIGQOOTSIRHAQIOLISPAVSASTNGREKICKADLEEV 432
Db 633 LIGPRFAEK--STYEFVGTSLDEDTSL-----PEGLKDMKNERK-----EKEEP 679
OY 433 SG 434
Db 680 PG 681

```

```

RESULT 7
US-08-928-213B-11
; Sequence 11, Application US/08928213B
; Patent No. 6238905
GENERAL INFORMATION:
APPLICANT: McHenry, Charles S.
Seville, Mark
Cull, Millard G.
TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III
HOLENZYME
NUMBER OF SEQUENCES: 195
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: CA
COUNTRY: USA

```



RESULT 9  
 US-09-388-774-1  
 Sequence 1, Application US/09388774  
 Patent No. 6228991  
 GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Guebler, Karl J.  
 APPLICANT: Patterson, Chandra  
 TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE  
 TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR

NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/388,774  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/074,579  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cerrone, Michael C  
 REGISTRATION NUMBER: 39,132  
 REFERENCE/DOCKET NUMBER: PF-0505 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 942 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: UTRSNOT02  
 CLONE: 688183  
 US-09-388-774-1

Query Match 4.6%; Score 104.5; DB 4; Length 942;  
 Best Local Similarity 18.0%; Pred. No. 0.41;  
 Matches 87; Conservative 70; Mismatches 132; Indels 195; Gaps 23;  
 93 PCPMWSEYSSSEVKTEYLMENFRALGRIKENK-----EYEGCVTLSP 140  
 98 FTMLIGDKYOGELTERE-----KKGDRVKEKRNKTEENGEKTEIFRASAVIPSK 150  
 141 EEA-----ESTGGYAKSISHYIISLKYKGTOKLDDSSIVDA----- 179  
 151 DKAAFLSYEELORRLKYEHS-----ISVRPOOLSGRLSDVNILLESAGIASLEVLPL 205  
 180 -----LIKEVAVGVITY--TEANSNAVRRGRDSDATEYDE 216  
 206 HNSRORGSGEDDSDPPPTVINONETFANIIIFKPTVVOQARIADONGLIDFIIINYDYN 265  
 217 AEEVY-----PIPKG-----EYHKKKEIVODVTLHD 242  
 266 REQSIGIOLVINGFYVFAFKDLPLPKKNVFEVLDSASWGTAKLRQKDALFTI-LHD 324  
 243 LDAANAPOGODILSLMG-----QMMKPRKTEITER 274  
 325 L-----RP-----QDRFSLIGSNRIKWKDHLISVTPDSIRDKVYIHHMSPGTGIDINCA 376  
 275 LROINKVNVKRI-DEGIAE-----LVPGVLFIDEVHMLDIECFSYLNRALESPLSP 325  
 377 LQRAI-RLNKKVYAHSGIGDRSVSLIVFLDCKPTVGEHTHTLKI-----LNTREARNGO 430  
 326 IYIATNRGICNVGTDMSPHICIPVDLDRLVYITRTETGPTENIQILAIRAOVEETDM 385  
 431 VCIFFTIGI-----NDVDRLEKLTSL-ENCGITRHHV-----EEDDA 467  
 386 DEESLAVLGEICGQSTLRHAIOI-ISPASVSVSKT-----NGREKICKADLEEVSGLY 436

Db 468 GSOLICFYDEL--RPLSLDIRIDYPPSSVQATKTLFPNFNSSEII-----IAGKL 518  
 QY 437 LPAK 440  
 Db 519 VDRK 522

RESULT 10  
 US-08-928-213B-10  
 ; Sequence 10, Application US/08928213B  
 ; Patent No. 6238905  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McHenry, Charles S.  
 ; Seville, Mark  
 ; Cull, Millard G.  
 ; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III  
 ; HOLOENZYME  
 ; NUMBER OF SEQUENCES: 195  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MEDLEN & CARROLL, LLP  
 ; STREET: 220 Montgomery Street, Suite 2200  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,213B  
 ; FILING DATE: 12-Sep-1997  
 ; CLASSIFICATION: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MacKnight, Kamrin T.  
 ; REGISTRATION NUMBER: 38,230  
 ; REFERENCE/DOCKET NUMBER: ENZYCO-02550  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-705-8410  
 ; TELEFAX: 415-397-8338  
 ; INFORMATION FOR SEQ ID NO: 10:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 273 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
 US-08-928-213B-10

Query Match 4.5%; Score 101.5; DB 4; Length 273;  
 Best Local Similarity 25.6%; Pred. No. 0.11;  
 Matches 55; Conservative 24; Mismatches 73; Indels 63; Gaps 10;  
 224 PKGEVHKKKEIYVD--VTLDDLAANAPOGODIISLMGQMMKPRKTEITERLROEINK 281  
 79 PCGVCDNCREIEGRFVYDLLEIDASR-----TKVEDTRDLIDNV-- 118  
 282 VVNRVIDEGIAELVPG-----VLFIDEVHMLDIECFSY-LNRALESPLSP---IYIATN 332  
 119 -----QYAPARGRFKYVLIDEVHMLSRHSFNALLKTLDEPXXPEHVKFIAT- 166  
 333 RGICNVGTDMSPHICIPVDLDRLVYITRTETGPTENIQILAIRAOVEETDMDESLAY 392  
 167 -----TDPOKLPVTILSR--CLQFHXHXXLKALDVEQIRHOLEHI-LNEEHIAH 212  
 393 LGEIQ-----QTSLRHAIOIISPAVSVSKTNGR 421  
 213 EPRAQLIARAAGSLRDALSLTDQATASGXDDQ 247

```

RESULT 11
US-09-541-782-2
; Sequence 2, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Berand, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1184
; TYPE: PRF
; ORGANISM: A. nidulans
US-09-541-782-2

```

```

Query Match 4.4%; Score 100.5; DB 4; Length 1184;
Best Local Similarity 18.3%; Pred. No. 1.4;
Matches 75; Conservative 81; Mismatches 167; Indels 87; Gaps 17;

```

```

OY 86 ELGSKVPPCPVGVSEVSVSEVKTVEVLMENFRRAIGRIKENEV-YEGEVELSPEAE 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 QINSTPMPTLREFTAIEIKLKAELLATRRHRCVYVSEYEMKRNESRIISEQR 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 145 STTGAYAKSHVILSKTVKGT-KOLKDSIYDALIKEKAVAGVI---YIANGSA 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 AIESMESIRKIVQELLITFSKFNLDKKN---DPTLAALCSTNDVLOOTDVLONTA 537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 200 ---VRVGCDFSEATYDLAEVYPIPKGVHKKELVODVT---LHDLDAANA 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 538 QLEEBMLCAHEETEHOIQ---DVGKGLITLGVGVEDINSLOSKLDRAKALDFTNA 592
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 249 QVGGODLISLMGMMKPKTEITEKLRQELINKVRY-----IDEGIAELVPGV 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 593 E-----LMRASSTVEDVTKRIDQRYEAFQTRAKLLETTSVAVNEFIATEISNI 642
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 299 LFIDEVHMLDIECFSYLNALFESPLSPYIATNNGICNVAGTDMTSPHGIPVLDRLV 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 643 ---ETRSDL---SEYNSISDA-----ACNNAKAETSSAHEDMNVLEIK 682
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 359 TIRTEYTGPT--EMIOILAIRAOVEIDMDESLAYLGEIGQTSIRNA-----IOL 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 683 DLKEEVKSKVGEGLNGLSAAARISF-----EVIGEPQLHLSQLTSPNLSGKDLKS 734
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 409 I--SPASVSKTNGREKICKADLEEVSGLYDA--KSSARLLQEQOEVI 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 735 IFETMATHLSECKNEINRLRAELQSSNRONITFTHKASAHLAQALEEHV 784
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
US-08-466-390-4
; Sequence 4, Application US/08466390
; Patent No. 5686562
; GENERAL INFORMATION:
; APPLICANT: TOURKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,390
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTD-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7100
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-390-4

```

```

Query Match 4.3%; Score 98; DB 1; Length 2101;
Best Local Similarity 20.7%; Pred. No. 6.1;
Matches 107; Conservative 88; Mismatches 190; Indels 132; Gaps 28;

```

```

OY 4 EEVOSTKQRIATHTHIKGIGLDANGMALAAGFYGOAARE-----AAGL 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1340 EQALSTLOLEHTSTQALVSEL-LPAKHLCOOLAQ---EQAAAEKRHRELEQSKQAAGGI 1395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 52 AVDMRQKKMRAVLLAGPATGKTALALGIAQEL-GSVVPPCPVGVSEVSEVKTVE 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1396 RAELRAQRELCGLI-----PLRQKVAEQERTAQOULAERKASYAEOL---SMLKAH 1444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 111 VLMEFRRAIGRIKENEVYGE-----VIELS-----PEAESEST--- 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1445 GLAEENRGLIGERANLQGFLEVLDAQREKYVOELAAVRADAEFRLAEVQREKQSTARE 1504
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 147 ---TGGYAKSHVILSKTVKGT-KOLKDSIYDALIKEKAVAGVIYIANGSAVKR 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1505 LEVMTAKTEGAAVVKVLEERQREERQ-KITAOVEE--LSKKLSDS---QASKVQOOK 1557
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 203 VGRCSFATEYDLAEVY---VPIPKGVHKKELVODVTILH-----DIDAANAQVGG 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1558 IKAVQAQGESEGOEORFOAQLNELQAOLSOKEAAEHYKLOMERAKTHYDAKQAOQEL 1617
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 254 QDILISLMGMMKPKTEI---TEKLRQELINKV-----VNRVIDEGIAELVPGVLET 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1618 QEOQLRSLQLOLQENK-ELAAEAERLGHLEQAGLTKTEAEQTCRHLLTAQVRSLEAOVNA 1676
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 302 DEVHMLDIECFSYLNALFES--PLSPYIATNNGICNVAGTDMTSPHGIPVLDRLV 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1677 DQ-QLRDLQKFQVATDALASREPOAKPQIDL-----SIDSLDSCBEGTPLSTSKLP 1728
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 359 TIRTEYTG---PEMIOILAIR--AOVEIDM-----DEESLAYLGEI--- 396
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1729 --RTQPDGTSVGPSPASPIQSRLPKVSELSIYFTPIPARSQAPLESLSLDSLGVPIDS 1786
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 397 GQOCT--SLRHAQILSPASVSKTNGREKICKADLEE 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1787 GRKTRSAARRTQILN---ITWTK-----KIDVBE 1813
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
US-08-470-950-4
; Sequence 4, Application US/08470950
; Patent No. 5698439
; GENERAL INFORMATION:
; APPLICANT: TOURKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE

```

```

TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,950
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7100
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-950-4

```

```

Query Match      4.3%; Score 98; DB 1; Length 2101;
Best Local Similarity 20.7%; Pred. No. 6.1;
Matches 107; Conservative 88; Mismatches 190; Indels 132; Gaps 28;

```

```

QY 4 EVOVSTSKKRIATHTHIGKGLDANGMAIALAGFVGQAARE-----AAGL 51
DB 1340 EQALSTLLEHTSTQALVSEL-LPAKHLCQQLA---EQAAEKRRHELEDSKQAAAGL 1395
QY 52 AVDMIRKMKAGRAVLLAGPRATGKTALALGIAQEL-GSKVPCPMVGSSEYVKTE 110
DB 1396 RAELRAQRELGLI-----PLRKYAEQERTAOQLRAEKASYAEOL-----SMKKAH 1444
QY 111 VLMEFRRAIGLRIKENKEVEGE-----VTELS-----PEAEEST--- 146
DB 1445 GLIAEENRGLGERANLGRQFLVELEDOAREKYVDLAAVRADETRLAEOVREAOSTARE 1504
QY 147 ----TGGAKSISHYIISLKTAVGTQOLKIDSSIYDALIKEKVAAGDVITYEANGAVKR 202
DB 1505 LEVMTAKTEGAAVKVLEERQFQERO-KLTAQVEE--LSKKLADSD---QASKVQOQK 1557
QY 203 VGRCDSPATEYDLEAEFY--VPIPKGEVHKKEIYQDVTLL-----DLDAANQPOGG 253
DB 1558 LKAVOAGGSGEQEQRQAOLNELQALQSOKEQAHEHYKLOMEKAKTHYDAKKQONQEL 1617
QY 254 ODILSLMGQMMKPKRTET---TEKLROEINKV-----VNRVIDEGIAELVPGVLFI 301
DB 1618 QEOURLSLEOLQENK-ELRAEAERLGHLEQAGLKTKEAQCRLHTLAQVRSLEAOVANA 1676
QY 302 DEVHMLDIECFSYLNRALS---PLSPIVILATNRGICNVGCDMTSPHGIPVLLDRLV 358
DB 1677 DQ-QLRDLGKQVATDLAKSHEPOAKPOLDL-----SLSIDLSCEGEGPLISTSKLP 1728
QY 359 IIRREYVG--PTEMIOILAIR--AQVEEIDM-----DESLATLGLF----- 396
DB 1729 --RTOPGTSVPGEPASPIQORLPKVESLESILYTPIPARSCAPLESSLSDIGOVFLDS 1786
QY 397 GQOQ--SLRAIALQILASPAVSKTNGREKICADLEE 431
DB 1787 GRKTRSAKRRRTQIIN---ITWTK-----KLDAVEE 1813

```

```

RESULT 14
US-08-467-781-4
Sequence No. 5780596
Patent No. 5780596

```

```

GENERAL INFORMATION:
APPLICANT: TOURKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,781
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7100
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-781-4

```

```

Query Match      4.3%; Score 98; DB 1; Length 2101;
Best Local Similarity 20.7%; Pred. No. 6.1;
Matches 107; Conservative 88; Mismatches 190; Indels 132; Gaps 28;

```

```

QY 4 EVOVSTSKKRIATHTHIGKGLDANGMAIALAGFVGQAARE-----AAGL 51
DB 1340 EQALSTLLEHTSTQALVSEL-LPAKHLCQQLA---EQAAEKRRHELEDSKQAAAGL 1395
QY 52 AVDMIRKMKAGRAVLLAGPRATGKTALALGIAQEL-GSKVPCPMVGSSEYVKTE 110
DB 1396 RAELRAQRELGLI-----PLRKYAEQERTAOQLRAEKASYAEOL-----SMKKAH 1444
QY 111 VLMEFRRAIGLRIKENKEVEGE-----VTELS-----PEAEEST--- 146
DB 1445 GLIAEENRGLGERANLGRQFLVELEDOAREKYVDLAAVRADETRLAEOVREAOSTARE 1504
QY 147 ----TGGAKSISHYIISLKTAVGTQOLKIDSSIYDALIKEKVAAGDVITYEANGAVKR 202
DB 1505 LEVMTAKTEGAAVKVLEERQFQERO-KLTAQVEE--LSKKLADSD---QASKVQOQK 1557
QY 203 VGRCDSPATEYDLEAEFY--VPIPKGEVHKKEIYQDVTLL-----DLDAANQPOGG 253
DB 1558 LKAVOAGGSGEQEQRQAOLNELQALQSOKEQAHEHYKLOMEKAKTHYDAKKQONQEL 1617
QY 254 ODILSLMGQMMKPKRTET---TEKLROEINKV-----VNRVIDEGIAELVPGVLFI 301
DB 1618 QEOURLSLEOLQENK-ELRAEAERLGHLEQAGLKTKEAQCRLHTLAQVRSLEAOVANA 1676
QY 302 DEVHMLDIECFSYLNRALS---PLSPIVILATNRGICNVGCDMTSPHGIPVLLDRLV 358

```

```

Db 1677 DQ-QLRDLGKFOVATDALKSREPQAKPOLDL-----SIDSLDSCBEGTPLSTSKLP 1728
QY 359 IIRRETYG---PTEMIQILAIR--AOVEIDM-----DESLAYIGE1----- 396
Db 1729 --RTQPDGTSVGPBPASPIQSRLPPKVESLESYFTPIPARSQAPLESIDSLDGVFLDS 1786
QY 397 GOOT--SLRHA1QILISPAVSCKTNGREKICKADLEE 431
Db 1787 GRKTRSAARRRTQIIN-----ITMK-----KLDVEE 1813

```

```

RESULT 15
US-08-483-924-4
; Sequence 4, Application US/08483924
; Patent No. 5882876
; GENERAL INFORMATION:
; APPLICANT: TOOKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,924
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-483-924-4

```

Query Match 4.3%; Score 98; DB 2; Length 2101;  
 Best Local Similarity 20.7%; Pred. No. 6.1;  
 Matches 107; Conservative 88; Mismatches 190; Indels 132; Gaps 28;

```

QY 4 EEVOSTSKKRIATHTHIKGLIDANGMAIALAAGFVGQAARE-----AAGL 51
Db 1340 EQALSTLOLEHTSTQALVSEL-LPAKHLCQQLQA---EQAAAEKRRRELDQSKQAAGL 1395
QY 52 AVDMIROKKMGRAVLLGPPATGKTALALGIADEL-GSKVPCPMGSEVYSSEVKTE 110
Db 1396 RAELLRAQRELELTI-----PLROKVAEQERTAOQLRAEKASYAEOL-----SWLKRAH 1444
QY 111 VMENFRRAIGLRIRKENKEVEGE-----VPELS-----PEEAEST--- 146
Db 1445 GLLAEENGIGLRANGLQFLEVEDQAREKYVOELAAVRADATRLAEVQREAOSTARE 1504
QY 147 ---TGYAKSISHVITSLKTVKTKQLKLDSSIYDALIKEKVAAGVDITYEANGAVKR 202
Db 1505 LEVMTAKYEGAAVVKLEERQRFQERQ-KLTAQYEE--LSKKIADSD---QASKVQOQK 1557

```

```

QY 203 VGRCDSPATREYDLAEAY---VPIRGEVHKKEIYQDVTIL-----DLDAANAQPPGG 253
Db 1558 LKAVQAQGGESQOEAORFOAQLNELQAQLSOKEOAAEHYKLMKEKAKTHYDAKKQOQNEL 1617
QY 254 ODLISLMGOMKPRKTEI---TEKLROEINKV-----VNR1DEGIAELVGVGLFI 301
Db 1618 QEOQLRSLEBOLQENK-ELRAEAERLGHLELOAGLKTKEAQOTCRHLTAQVRSLEAQVANA 1676
QY 302 DEVMHLDIECFSTYLNRALES---PLSPIVILATNRGICVNRGDMTSPHIGIVDLDRLY 358
Db 1677 DQ-QLRDLGKFOVATDALKSREPQAKPOLDL-----SIDSLDSCBEGTPLSTSKLP 1728
QY 359 IIRRETYG---PTEMIQILAIR--AOVEIDM-----DESLAYIGE1----- 396
Db 1729 --RTQPDGTSVGPBPASPIQSRLPPKVESLESYFTPIPARSQAPLESIDSLDGVFLDS 1786
QY 397 GOOT--SLRHA1QILISPAVSCKTNGREKICKADLEE 431
Db 1787 GRKTRSAARRRTQIIN-----ITMK-----KLDVEE 1813

```

Search completed: November 13, 2001, 07:01:18  
 Job time: 324 sec





## alignment\_block:

US-09-589-510-4 x AAD02567

Align seg 1/1 to: AAD02567 from: 1 to: 1912

```

1 MetArg1IleGluIuValGlnSerThrSerLysLysGlnArg1IleAlaThr 17
  |||
94 ATGAGATCGAGGAGGTGCAGATCGACCTCGAAGACGACGACATCGGCAC 143
  |||
17 RHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
  |||
144 CCACACCCACATCAAGGAGACTCGGCTCGACGCCAATGGATGGCCGATTG 193
  |||
34 IAlaAlaIleAlaGlyPheValGlyGlnAlaIleAlaArgGluAlaIleA 50
  |||
194 CGTTGGGGGGGGGTTCTGTGGGGCCGAGCGGGCGGCGGAGGGCGGG 243
  |||
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaIleArgAlaValLe 67
  |||
244 CTGGCGGTGCACATGATTCGCCAGAGAGAGATGGCCGCGCGGCGGTGCT 293
  |||
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
  |||
294 CTTGGCGGGTCCGCCGCCAGGGCAAGCGCGCTAGCGCTCGGCATAG 343
  |||
84 IAlnGlnIleuGlySerLysValProPheCysPrometValIAlnSerGlu 100
  |||
344 CCCAGAGACCTCGGACGACAGAGTCCCTTCTGTCTATGTAAGTACGAA 393
  |||
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
  |||
394 GTGTACTCTCTCGAGGTCAAGAAAACGAGTGTGATGGAATAATTCCG 443
  |||
117 gArgAlaIleGlyLeuArgGlyIleLysGluAsnLysGluValTyrGluGly 134
  |||
444 TAGAGCTATAGGTTGGCTATAAAGAAAACAAAGAGTTATGAGGAG 493
  |||
134 IuValThrGluLeuSerProGluIuValGlnSerThrThrIleGlyTyr 150
  |||
494 AGGTTACTGAACTTCCCGAAGAGAGCTGAGAGTACAACTGAGTAT 543
  |||
151 AAlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyThr 167
  |||
544 GCAAAAAGCATTAACCATTAATCATCAGCTTAAGACTGTAAAGGAGAC 593
  |||
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
  |||
594 TAAAGCACTGAAGTTAGATTCTTCAATTATGATGCTGTGATCAAGGAAA 643
  |||
184 ySValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
  |||
644 AGGTGGGAGGTGGGTGATGTTATATACATGCAACAAATAGTGCAGCAG 693
  |||
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217
  |||
694 AAAAGAGTTGCTGATGCTGATCTTTTGGTACAGAAATAGCATCTTGAAG 743
  |||
217 aGluGluTyrValProIleProLysGlyGluValHisLysLysLysGlu 234
  |||
744 TGAAGAGTATGTTCTTATCCCAAGGTGAAGTCCATAGAAAAGAAA 793
  |||
234 IAlaValGlnAspValThrLeuHisAspLeuAspAlaIleAlaAsnIleAla 250
  |||
794 TTGTGCGAGGTGTCACATTCATGACCTTGATGTCACAAATGCTCAGCCA 843
  |||
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLeuLysProAr 267
  |||
844 CAAGGTGGCCAGATATTTGTCCCTTATGGGCGCAGATATGTAACACAG 893
  |||
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
  |||
894 AAAGACTGAATCACCGAAAACCTAGCCAGAAATTAATTAAGTGTGTA 943
  |||

```

```

284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
  |||
944 ATGATATATTCGATGAAGCAATTCGAGACTGTACCTGCTGTGTTGTTTC 993
  |||
301 IleAspGluValHisMetLeuAspIleGlyCysPheSerTyrLeuAsnAr 317
  |||
994 ATTGATGAGGTCCACATGTTGGATATCGAATGTTTCTTATCTTAACCG 1043
  |||
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
  |||
1044 TGCATTGGAGAGCCCATTTATCCCAATCTGTACTTCTTACAAATTAAGG 1093
  |||
334 IYIleCysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350
  |||
1094 GAAATATGTAATGTAAGAGCACTGATATGACAGATGCCATGTGTATACC 1143
  |||
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPr 367
  |||
1144 GTGGATCTTCTAGATAGGCTGTGATTTATTCGACAGAGACATATGGCCC 1193
  |||
367 oThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGluIleA 384
  |||
1194 TACTGAGATGATACAGATATTGCTATCCGACACAGAGTGGAGGAGATTG 1243
  |||
384 sPMeTAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGlnIleThr 400
  |||
1244 ATATGATGAAGAAAGCTCTTGCTTATTTAGCGAGATGGACAGACAGCA 1293
  |||
401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValSerLys 417
  |||
1294 TCTTTGAGACATGCTATTCATTTGATATCACCTGCCACGCGTCTCAA 1343
  |||
417 sThrAsnGlyArgGlyLysIleCysLysAlaAspLeuLysGluValSerG 434
  |||
1344 GACTATATGGAAGAGAAAATCTGCAGAGGCTGATCTCAGGAGACAGCTG 1393
  |||
434 IYLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlnGln 450
  |||
1394 GGCCTATTGATGATGCCAAATCTCGGCTCGGCTGCTCAGAGACAGCA 1443
  |||
451 GluArgTyrIleThr 455
  |||
1444 GAAGATACATCACC 1458
  |||
seq_name: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT.AAD02566
seq_documentation_block:
ID AAD02566 standard; cdNA; 1845 BP.
XX
AC AAD02566;
XX
DT 02-MAY-2001 (first entry)
XX
DE Maize RuVB orthologue #1 cDNA.
XX
KW Maize; RuVB orthologue; branch migration; heteroduplex extension;
homologous recombination; transformation; transgenic plant; ss.
XX
OS Zea mays.
XX
FT Key Location/Qualifiers
FT CDS /*tag= a
/*product= "Maize RuVB orthologue protein #1"
XX
PD 25-JAN-2001.
XX
PN
XX
XX
PF 13-JUN-2000; 2000MO-US16271.
XX
XX 16-JUL-1999; 99US-0144112.

```



XX PA (PION-) PIONEER HI-BRED INT. INC.  
 XX PI Mahajan PB;  
 XX PI WPI; 2001-159537/16.  
 DR P-PSDB; AAY72562.  
 XX Novel maize Ruvb nucleic acid useful for modulating levels of maize  
 PT Ruvb in plants, as probes or amplification primers in the detection,  
 PT quantitation or isolation of gene transcripts  
 PS Claim 1; Page 67-69; 87pp; English.  
 CC The present sequence is a Zea mays Ruvb orthologue #1 cDNA. Ruvb along  
 CC with Ruvb catalyses the branch migration process, also known as  
 CC heteroduplex extension, in homologous recombination. Ruvb is used for  
 CC the control of homologous recombination or transformation efficiency in  
 CC transgenic plants. The Ruvb nucleotide may be used as probes or  
 CC amplification primers for detecting, quantifying or isolating gene  
 CC transcripts, in detecting deficiencies in the mRNA level during screening  
 CC for desired transgenic plants, for detecting gene mutations, for  
 CC monitoring upregulation of expression or changes in enzyme activity, for  
 CC detecting any number of allelic variants, orthologues or paralogues of  
 CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
 CC also be used for recombinant expression of its encoded polypeptide, or  
 CC for use as immunogen in preparing and/or screening of antibodies, and in  
 CC sense or antisense suppression of one or more genes in a host cell,  
 CC tissue or plant. The Ruvb proteins may be used in assays to agonise or  
 CC antagonise the enzyme function, or as immunogens or antigens for  
 CC screening antibodies.  
 SO Sequence 1845 BP; 530 A; 378 C; 490 G; 445 T; 2 other;

alignment\_scores:  
 Quality: 2253.00 Length: 455  
 Ratio: 4.952 Gaps: 0  
 Percent Simlarity: 100.000 Percent Identity: 99.341

alignment block:  
 US-09-589-510-4 x AAD02566 ..

Align seq 1/1 to: AAD02566 from: 1 to: 1845

117 GARGAlaIleGlyLeuAArgTlleGlySlnuAsnLysGluValTyrGluGlyG 134  
 435 TAGAGCTATAGGTTTGGCTATAAAGAAAAACAAAGAGCTTATGAAGAG 484  
 134 LuValThrGluLeuSerProGluGluAgluSerThrThrGlyGlyTyr 150  
 485 AGGTTACTGAACCTTCCCGACAGAGAGCTGAGATACACCTGGGAGATAT 534  
 151 AlalysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167  
 535 GCAAAAGACATTAGCCATGATATCATCGCTTAAAGACTGTAAAGAGAC 584  
 167 rlyGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGluL 184  
 585 TAGAGCACTGAAGTTAGATCTTCAATTTATGATGCTGATCAAGGAAA 634  
 184 ysaValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200  
 635 AGGTGGCAGTGGTGATCTTATATACATTGAAGCAAAATGAGCAGATG 684  
 201 LysArgValAlGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
 685 AAAAGAGTTGGTAGATGTGATTTCTTTCTACAGAAATACGATCTTGAAGC 734  
 735 TGAGAGATATGTTCATCCCAAGCCCAAGGTGAAGTCCATGAAGAAAAAGAA 784  
 217 aGluGluTyrValProIleProLysGlyGluValHisLysLysGluT 234  
 785 TAGTGAGAGATGTCAACACTTCACTGACCTTGATGACGCAAAAGCTCAGCCA 834  
 234 lvalGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGluPro 250  
 251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267  
 835 CAAGGTGGCCAGATATTTTGTCCCTTANGGGCCAGATGATGAACACAG 884  
 267 GlyThrGlnIleThrGluLysLeuArgGlnIleLysLysValAla 284  
 885 AAGACTGAATATCCCGAAAAACCTACGCCAAGAAATTAATAAGGGGTA 934  
 284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300  
 935 ATGATATATCGATGAAGAAATTCAGAGCTTGATCTGTGTTTATTC 984  
 301 lIleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317  
 985 ATTGACAGAGGTCCACATGTTGATATCGAATGTTTTCTTATCTTAACCG 1034  
 317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334  
 1035 TGCATTGAGAGCCCATTTATCACCATTGTGATACTTGCTACGAATAGGG 1084  
 334 lylIecysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350  
 1085 GAATATATTAATGTAAGAGAACCTGATATGACAAGTCCACATGTTANACA 1134  
 351 ValAspLeuLeuAspArgLeuValIleIleIleArgThrGluThrTyrGlyP 367  
 1135 GTGATCTTCTAGATAGCTGTGATGATTTATTCGAGACAGAACATATGCC 1184  
 367 cThGluMetIleGlnIleLeuAlaIleArgAlaGlnValAlaGluIleAla 384  
 1185 TACATGAGATGATACAGATATATGGCTATCCAGACCAAGTGGAGACATTG 1234  
 384 spMeLAspGluGluSerLeuAlaTyrLeuGlyIleGlyGlnIleThr 400  
 1235 ATATGAGATGAAGAAAGCTTGTATATTAGCGAGATCGGACAGCAGACA 1284  
 401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerL 417  
 1285 TCTTTAAGACATCTATTTCAATTGATATCACTGCGCAGCGCTGCTCAAA 1334  
 417 sThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluGluValSerG 434

1335 GACATAAGCAGAGAAAATATGCAGGCTATCTCAGGAGGTCATG 1384  
434 IYleuYrIleuSpAlaYsSerSerAlaArgLeuIleuGlnGlnGln 450  
1385 GGCTCTATTGGATGCGCAAAATCCTCGCGCTCTCCAGAGCAACAA 1434  
451 GluArgTyrIleThr 455  
1435 GAAAGATACATCACCC 1449

seq\_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD02568

seq\_documentation\_block:

ID	standard; cdna; 1886 BP
AD02368	
AD02369	
AD02370	
AD02371	
AD02372	
AD02373	
AD02374	
AD02375	
AD02376	
AD02377	
AD02378	
AD02379	
AD02380	
AD02381	
AD02382	
AD02383	
AD02384	
AD02385	
AD02386	
AD02387	
AD02388	
AD02389	
AD02390	
AD02391	
AD02392	
AD02393	
AD02394	
AD02395	
AD02396	
AD02397	
AD02398	
AD02399	
AD02400	
AD02401	
AD02402	
AD02403	
AD02404	
AD02405	
AD02406	
AD02407	
AD02408	
AD02409	
AD02410	
AD02411	
AD02412	
AD02413	
AD02414	
AD02415	
AD02416	
AD02417	
AD02418	
AD02419	
AD02420	
AD02421	
AD02422	
AD02423	
AD02424	
AD02425	
AD02426	
AD02427	
AD02428	
AD02429	
AD02430	
AD02431	
AD02432	
AD02433	
AD02434	
AD02435	
AD02436	
AD02437	
AD02438	
AD02439	
AD02440	
AD02441	
AD02442	
AD02443	
AD02444	
AD02445	
AD02446	
AD02447	
AD02448	
AD02449	
AD02450	
AD02451	
AD02452	
AD02453	
AD02454	
AD02455	
AD02456	
AD02457	
AD02458	
AD02459	
AD02460	
AD02461	
AD02462	
AD02463	
AD02464	
AD02465	
AD02466	
AD02467	
AD02468	
AD02469	
AD02470	
AD02471	
AD02472	
AD02473	
AD02474	
AD02475	
AD02476	
AD02477	
AD02478	
AD02479	
AD02480	
AD02481	
AD02482	
AD02483	
AD02484	
AD02485	
AD02486	
AD02487	
AD02488	
AD02489	
AD02490	
AD02491	
AD02492	
AD02493	
AD02494	
AD02495	
AD02496	
AD02497	
AD02498	
AD02499	
AD02500	
AD02501	
AD02502	
AD02503	
AD02504	
AD02505	
AD02506	
AD02507	
AD02508	
AD02509	
AD02510	
AD02511	
AD02512	
AD02513	
AD02514	
AD02515	
AD02516	
AD02517	
AD02518	
AD02519	
AD02520	
AD02521	
AD02522	
AD02523	
AD02524	
AD02525	
AD02526	
AD02527	
AD02528	
AD02529	
AD02530	
AD02531	
AD02532	
AD02533	
AD02534	
AD02535	
AD02536	

AC AAD02568;

02-MAY-2001 (first entry)

DE Maize Ru5B orthologue #3 cDNA.  
XX

XX homologous recombination; transgenic plant; ss.

Ph	key	Location/Qualifiers
FT	CDS	82..1449
FT		/*tag= a
XX		/product= "Maize RuvB orthologue protein #3"

PN WO200105975-A1  
XX

PD 25-JAN-2001

13-JUN-2000; 2000WO-US16271.

PR 16-JUL-1999; 99US-0144112.  
XX

PA (PLON-) PIONEER HI-BRED INT INC.  
XY

PI Mahajan PB;

DR WPI; 2001-159537/16.  
DR D-PCDD; 19970554

DR P-PSDB; AAY72564.

quantitation or isolation of gene transcripts -

Disclosure; Page 74-76; 87pp; English.

The present sequence is a *zea mays* RuvB orthologue #3 cDNA. RuvB along with RuvA catalyses the branch migration process, also known as heteroduplex extension, in homologous recombination. RuvB is used for the control of homologous recombination or transformation efficiency in transgenic plants. The RuvB nucleotide may be used as probes or amplification primers for detecting, quantifying or isolating gene transcripts. In detecting deficiencies in the mRNA level during gene transfer for desired transgenic plants, for detecting gene mutations, for monitoring upregulation of expression or changes in enzyme activity, for detecting any number of allelic variants, orthologues or paralogues of the gene, or for site directed mutagenesis in eukaryotic cells. It may also be used for recombinant expression of its encoded polypeptide, or for use as immunogen in preparing and/or screening of antibodies, or sense or antisense suppression of one or more genes in a host cell, tissue or plant. The RuvB proteins may be used in assays to agonise or antagonise the enzyme function, or as immunogens or antigens for screening antibodies.

Sequence 1886 BP; 524 A; 408 C; 518 G; 436 T; 0 other;

```
alignment_scores:
```

Quality:	2222.00	Length:	455
Ratio:	4.916	Gaps:	0
Percent Similarity:	99.341	Percent Identity:	97.582

```
alignment_block;
```

US-09-589-510-4 X AAD02568

Align seg 1/1 to: AAD02568 From: 1 to: 1886

[illegible]

```

832 CAAGGTGGCCAAAGATATTTTGTCCCTTATGCGCCAGATGATGAAGCCAGC 881
267 gLysThrGluLeuThrGluLysLeuArgGlnGluIleAsnLysValValA 284
882 GAAGACTGAATACCCGAAAGCTACGCCAGAAATCAATAGGTGGTAA 931
284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
932 ACAGTATATGACGAGAGGATCCAGAGCTTGACCGGTGGTTTGTTC 981
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
982 ATTGATGAGGTCCACATGTGGATATGAAAGCTTTCTTATCTTAACCG 1031
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
1032 TGCATTGGAGGCCCATTTATCCCAATTTGTATCTGCTACGCAATAGAG 1081
334 LylleCysAsnValAlaArgGlyThrAspMetThrSerProHisGlyIlePro 350
1082 GAATATGTAATGTGAGAGAACCGATATGACGATCCACATGATATCA 1131
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPr 367
1132 GTGGACCTTCTAGATAGGTGGTGTATATTCGACAGAAACATATGAGCCC 1181
367 cThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluLysLeu 384
1182 TACTGACATGATACGATACTGCTATCCAGACCAAGTGAAGAGATG 1231
384 spMetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGlnGlnThr 400
1232 ATATGATGAGAAAGAGTCTGCTATTATTAGCGAGATCGGACACACACA 1281
401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerly 417
1282 TCTTTGAACATGCTATTCAGTGTCTATCACTCCACGCGGTGGCAAA 1331
417 sThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluGluValSerG 434
1332 GACCAACGGGAGAGAAAGATGTCAAGGCTGACCTCGAGGAAGTCAAGCG 1381
434 lYleuTyrLeuAspAlaLysSerSerAlaArgLeuGlnGlnGln 450
1382 GGCTCTATTGATGCCAATCCTCGGCTCGCTCCAGACACACAA 1431
451 GluArgTyrIleThr 455
1432 GAAAGATACATCACCC 1446
seq_name: /SIDS1/gcgdata/geneseq/geneseqn/MA2001.DAT.AAD02569
seq_documentation_block:
ID AAD02569 standard; cDNA: 1898 BP.
XX
AC AAD02569;
XX
DT 02-MAY-2001 (first entry)
XX
DE Maize RuVB orthologue #4 cDNA.
XX
KM Maize: RuVB orthologue; branch migration; heteroduplex extension;
XX homologous recombination; transformation; transgenic plant; ss.
XX
OS Zea mays.
XX
Key Location/Qualifiers
FH 166..1536
CDS /*tag= a
FT /product= "Maize RuVB orthologue protein #4"
XX
XX WO200105975-A1.
XX

```

```

PD 25-JAN-2001.
XX
XX 13-JUN-2000; 2000WO-US16271.
XX
XX 16-JUL-1999; 99US-0144112.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPT: 2001-159537/16.
XX
XX P-PSDB: AAY72565.
XX
XX Novel maize RuVB nucleic acid useful for modulating levels of maize
XX RuVB in plants, as probes or amplification primers in the detection,
XX quantitation or isolation of gene transcripts -
XX
XX Disclosure: Page 77-79; 87pp; English.
XX
XX PS
XX
XX The present sequence is a Zea mays RuVB orthologue #4 cDNA. RuVB along
XX with RuVA catalyses the branch migration process, also known as
XX heteroduplex extension. In homologous recombination, RuVB is used for
XX the control of homologous recombination or transformation efficiency in
XX transgenic plants. The RuVB nucleotide may be used as probes or
XX amplification primers for detecting, quantifying or isolating gene
XX transcripts, in detecting deficiencies in the mRNA level during screening
XX for desired transgenic plants, for detecting gene mutations, for
XX monitoring upregulation of expression or changes in enzyme activity, for
XX detecting any number of allelic variants, orthologues or paralogues of
XX the gene, or for site directed mutagenesis in eukaryotic cells. It may
XX also be used for recombinant expression of its encoded polypeptide, or
XX for use as immunogen in preparing and/or screening of antibodies, and in
XX sense or antisense suppression of one or more genes in a host cell,
XX tissue or plant. The RuVB proteins may be used in assays to agonise or
XX antagonise the enzyme function, or as immunogens or antigens for
XX screening antibodies.
XX
XX SQ Sequence 1898 BP; 532 A; 413 C; 521 G; 432 T; 0 other;
XX
XX
XX alignment_scores:
XX Quality: 2211.50 Length: 456
XX Ratio: 4.893 Gaps: 1
XX Percent Similarity: 99.123 Percent Identity: 97.368
XX
XX alignment_block:
XX US-09-589-510-4 x AAD02569 ..
XX
XX Align seq 1/1 to: AAD02569 from: 1 to: 1898
XX
XX 1 MetArgIleGluGluValAlaGlnSerThrSerLysLysGlnArgIleAlaThr 17
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 166 ATGAGGATAGAGGAGGTGCAATCGACCTCGAAGAAACACACGCTGCCAC 215
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 17 rHisThrHisIleLysGlyLeuGlyLeuAsp...AlaAsnGlyMetAlaI 33
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 216 CCACACTCACATCAAGGCCCTCGCTCGACCAAGCGCATGCAATGTCCA 265
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 33 leAlaLeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaAla 49
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 266 TCGCGTTGGCGCGGGGTTCTGTGGCCAGGCGCGCGCGCGAGCGGCC 315
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 50 GlyLeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaVa 66
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 316 GGGCTGGCGGTGACATGATCCCGCAAGAAAGATGGCGCGTCCGCT 365
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 66 lleuLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyI 83
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 366 GCTCCTTGGCGGCGCGCGCCGCAAGCGGCAAAAGCGCTAGCGCTGGCA 415
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 83 leAlaGlnGluLeuGlySerLysValProPheCysProMetValGlySer 99
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 416 TACGCGAGGAGCTCGGCGACAGGTCCTCATCTGTCTATGTGATGATCA 465
XX

```



CC antagonise the enzyme function, or as immunogens or antigens for  
XX screening antibodies.  
SQ Sequence 1869 BP; 539 A; 373 C; 486 G; 471 T; 0 other;

alignment\_scores:  
Quality: 2157.50 Length: 455  
Ratio: 4.926 Gaps: 2  
Percent Similarity: 96.264 Percent Identity: 96.264

alignment\_block:  
US-09-589-510-4 x AAD02570 ..

Align seg 1/1 to: AAD02570 from: 1 to: 1869

```
1 MetArgIIeGIuGIuValGInSerThrSerLysLysGlnArgIIeAlaThr 17
|||||
64 ATGAGGATCGAGAGAGTGCAGTCGACCTCGAAGAAAGCAGCGCATGCCAC 113
17 rHsThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
|||||
114 CCACACCCCATCATCAGGAGCTGCGCTCGAGCGCAATGGGATGGCATTTG 163
34 lAlaValAlaIleGlyPheValGlyGlnAlaAlaAlaArgGlnAlaIleGly 50
|||||
164 CGTTGGCGGCGGGGTTCTGTTGGGCGGCGCGCGGCGGCGTGT 186
51 LeuAlaValAlaAspMetIleArgGlnLysLysMetAlaGlyArgAlaValIle 67
|||||
187 .....CAGAAGAAGATGGCGCGCGCGCGGCGGCGTGT 215
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
|||||
216 CCTTGGGGGTCGCCGCCACGGGAGAGAGCGGCTAGCG...GGCATGG 262
84 lArgGlnLysLeuGlySerLysValProPheCysProMetValGlySerGlu 100
|||||
263 CCAGAGAGCTCGGAGCAAGAGTCCCTTCTGCTATGATGATGATCAGAA 312
101 ValTyrSerSerGluValLysLysThrGluValIleMetGluAspPheAr 117
|||||
313 GTGTACTCTCGAGAGTCAAGAAACTGAGTGTGATGGAATAATTTCCG 362
117 gArGAlaIleGlyLeuArgIleLysGlnAsnLysGlnValTyrGlnGlyG 134
|||||
363 TAGAGCTATAGGTTGGTATTAAGAAACAAAGAGCTTTATGAGAGAG 412
134 luvAlThrGluLeuSerProGlnGluAlaGluSerThrThrGlyGlyTyr 150
|||||
413 AGGTACTGAACCTTCCCGAAGAGGCTGAGATACAACTGCTGATAT 462
151 AlAlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
|||||
463 GCAAAAAGCATTAAGCATGATATCATCAAGCTTAAGACTGTTAAAGGAG 512
167 rLysGlnLysLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
|||||
513 TAGGCACTGAGAGTATGATCTTCAATTATATGATCTCTGATCAAGGAAA 562
184 ySValAlaValAlaGlyAspValIleTyrIleGlnAlaAsnSerGlyAlaVal 200
|||||
563 AGGTGCGAGTGGTGTATGTTATATATCATCAAGCAAAATAGTGAGCAGTG 612
201 LysArgValAlaArgCysAspSerPheAlaThrGluTyrAspLeuGlnAl 217
|||||
613 AAAAGGTTGGTAGATGATCTTTTGTCTACAGATACGATCTTGAAGC 662
217 aGluGluTyrValProIleProLysGlyGlnValHisLysLysLysGluI 234
|||||
663 TGAAGAGTATGTTCTATCCCAAGAGTGAAGTCCATTAAGAAAAAGAAA 712
234 lValAlaGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
```

```
|||||
713 TTGTCCAGAGATGTCACACTTCATGACCTTGATGACGCAAAATGCTCAGCA 762
251 GInGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267
|||||
763 CAAGGTGGCCAAAGATATTTGTCCCTTATGCGCCAGATGATGAAGAACAG 812
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
|||||
813 AAAGACTAAATCACCCGAAAACTACGCGCAAGAAATTAATATAGTGCTAA 862
284 snArgTyrIleAspGluGlyIleAlaGlnLeuValProGlyValLeuPhe 300
|||||
863 ATAGATATATCGATCAAGCAATGAGAGCTTGATACCTGCTGTTTGTTC 912
301 lLeAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
|||||
913 ATTGATGAGGTCCACATGTTGATATCGAAATGTTTCTTAATCTTAACCG 962
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
|||||
963 TGCAATGAGAGCCCATTAATACCAATGATGATCTGCTACAAATAGGG 1012
334 lYlLeCysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350
|||||
1013 CAATATGTAATGTAAGAGAGACTGATATGACAGACTCCACATGATATACG 1062
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPr 367
|||||
1063 GTGGATCTTCTAGATAGAGCTGTGATATATTCGACAGACATATAGCCCC 1112
367 cThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGlnGluIleA 384
|||||
1113 TACTGAGATGATACGATATGCTGCTATCCGACAAAGTGAAGAGATTTG 1162
384 sPmetAspGluGluSerLeuAlaTyrLeuGlyGlnIleGlyGlnGlnThr 400
|||||
1163 ATATGATGAAAGAAAGTCTTCTTATTAGCGAGATCCGACACAGACA 1212
401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerLy 417
|||||
1213 TCTTTGAGACATGCTATTCATTCATTCATTCACCTCCGCGGTCTCAAA 1262
417 sThrAsnGlyArgGluLysIleCysLysAlaAspLeuGlnValValSerG 434
|||||
1263 GACTAATGGAAGAGAGAAATCTCAAGGCTGATCTCGAGAACTTACTG 1312
434 lYLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlnGln 450
|||||
1313 GGCCTATTTGATGCCAAATCTCGGCTCGGCTCGCTCCAGAGACACAA 1362
451 GluArgTyrIleThr 455
|||||
1363 GAAAGATACATCACCC 1377
seq_name: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC38992
seq_documentation_block:
ID AAC38992 standard; DNA; 1612 BP.
XX
XX AAC38992:
AC
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 22954.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
```

```
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145376.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149375.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
```

```

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157765.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment_scores:
  Quality: 1994.00      Length: 455
  Ratio: 4.542          Gaps: 0
  Percent Similarity: 96.484  Percent Identity: 85.055

```

```

alignment_block:
US-09-589-510-4 x AAC38992 ..

```

```

Align seg 1/1 to: AAC38992 from: 1 to: 1612

```

```

1 MetAglIEglUGlUValGlnSerThrSerLysLysGlnAtrGllEalA1Th 17
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
133 G7AAAGATTGAAGAAATACATCCACCGCTAAAGAAACAAACGATTGCTAC 182
17 HisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
193 TCACACCCATATCAAAAGCCTTGCGCTGCAGCCAACTGATCCCTATTA 232
34 IaLeuAlaIaGlyPheValGlyGlnAlaAlaAlaIaArgGluAlaIaGly 50
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
233 AATTGGCAGCTGATTTGTTGTCACAACTTGAGGCTAGACAGCAGCTTCGT 282
51 LeuAlaValaAspMetIleArgGlnLysLysMetAlaGlyArgAlaValE 67
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
283 CTGTGATTTGACATGATTAAACAGAAATGGCGGCAAGGCTCTTTT 332
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
333 GCTTCTGAGACTCTCTGGAAGCTGGAAACAGCTTTGCTCTTGGAATCT 382
84 IaGlnGluLeuGlySerLysValaProPheCysPProMetValaGlySerGlu 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
383 CTCAGAGCTGGAGCAAGGTTCCATCTCTCAATGTTGATCGAG 432

```

```

101 ValTyrSerSerGluValLysLysThrGluValLeuMetGlnAsnPheAr 117
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
433 GTTACTCATCAGAGGTTTAAAGAAACAGAGCTTCTATGAGCAATTTTGA 482
117 gArGAlaIleGlyLeuArgIleLysGlnAsnLysGluValTyrGluGlyG 134
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
483 ACGTGCCATTGGTCTACGTATCAAGGAACCAAGAAAGTCTATCAAGGGG 532
134 LuValThrGluLeuSerProGluGluAlaLysSerThrThrGlyLysTyr 150
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
533 AGGTCACCGAGCTGTCCAGCAAGAAACTGAAGCCCTCACTGAGGTTAT 582
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyH 167
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
583 GGTAAAGCATCAGCAGCTGTTGTAATTACCTCAAGACAGTCAAAGGAAC 632
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGluL 184
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
633 CAACATCTCAAAATTTGATCCCACTATCTATGATGCCCTTGATTAAGAAA 682
184 ysValAlaValaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
683 AGGTAGCTGTAGAGAGATGTATCTATATCGAAGCAAAACAGTCGAGCTGTC 732
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuVal 217
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
733 AAACGGGTAGTAGAGTAGATGATTTTGGCCACTGCAATTTGATTTGAGAG 782
217 aGluGluTyrValProIleProLysGlyGluValHisLysLysGlyLui 234
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
783 AGAAGATATGTTCCACTTCCCAAGAGAGAGTCCCAAAAAGAAAGAGAG 832
234 leValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnIaGlnPro 250
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
833 TACTGCAGGATGTCCACACTCCAAAGATCTGGATCCAGCAAAATGCTCAGCT 882
251 GlnGluGlyGlnAspIleLeuSerLeuMetGlyLysMetMetLysProAr 267
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
883 CAAGGTGGCCAGGATATACTTCTTTGATGGCCCAATATGATGAACCGG 932
267 gLysThrGluIleThrGlnLysLeuArgGlnGluIleAsnLysValaValA 284
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
933 GAAGACGAGATCATCGATATAGCTTCGGCAAGAAATTTAACAAAGTTGTGA 982
284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
983 ACCGATATATAGATGAAGGTGTGGCAGAGCTGTTCAGAGATTCTATTT 1032
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1033 ATTGATGAGGTTCTATATGCTTGTATATGAGTGTCTCTCATCTATGAACCG 1082
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1083 TGCTCTTGAGAGCCATATTATCTCCGATAGATATTGGCAACAATATAGAG 1132
334 LyIleCysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1133 GTCTTTGCAACGTAAAGGGAGCTGATATGCCAGCCCACTGAGAGTCCCT 1182
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyLys 367
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1183 ATTGATCATATTAGATCGATTGGTTATCATCCGAGCTCAAAATCTATATATCC 1232
367 oThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValaGluGluIleA 384
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1233 CTCGTAATATATCCAGATTATATAGCAATTCGTGGCAAGTTGAAGATTAA 1282
384 sPmetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGlnGlnThr 400
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1283 CCGTGGATGAAGAAATCTGGTTCTACTTGGGAGGATGTGGCCAAAGAACT 1332

```

```

401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerIy 417
|||||
1333 TCACTAAGCAGCGCTGTCAGCTTCTCTCTCCAGCATTTGTACGAA 1382
417 sTrAsnGlyArgGlnIlyIleCySLysAlaAspLeuGlnIlyValSerG 434
|||||
1383 AATCAATGGCGGTGACAAATATTGCAAGGCTGATATAGAGAGTAACAT 1432
434 lYleuTYrIleuAspAlaIySerSerAlaArgLeuGlnIlyGlnIly 450
|||||
1433 CACTCTACTTGGATGCTAAATCTTCACGAAGCTTTGCAATGACAAACA 1482
451 GluArgTYrIleThr 455
|||||
1483 GAAAAATACATCTCA 1497

seq_name: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT:AAx18199

seq_documentation_block:
ID AAX18199 standard; cDNA to mRNA; 1730 BP.
XX
AC AAX18199;
XX
DT 10-MAY-1999 (first entry)
XX
DE TIP49 coding sequence #2.
XX
KM TIP49; TATA binding protein; TBP binding protein;
DNA transcription control; ss.
XX
OS Rattus sp.
XX
PN W09900419-A1.
XX
PD 07-JAN-1999.
XX
PF 25-JUN-1998; 98WO-JP02836.
XX
PR 27-JUN-1997; 97JP-0187398.
XX
PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
PI Kishimoto T, Makino Y, Miwa S, Tamura T;
XX
DR WPI: 1999-095682/08.
XX
P-PSDB: AAW74417.
XX
PT TBP-binding protein with DNA helicase and ATPase activities - gene
PS encoding it, and antibodies recognising it.
XX
PS Claim 10; Page 43-45; 64pp; Japanese.
XX
CC This sequence encodes the TATA-binding protein (TBP) binding protein,
CC designated TIP49, of the invention. TIP49 and its associated nucleic
CC acids and antibodies are useful in investigation of the process of DNA
CC transcription control by TBP in vivo, and in detection of the blocking of
CC DNA transcription.
XX
SQ Sequence 1730 BP; 469 A; 401 C; 476 G; 384 T; 0 other:

alignment_scores:
Quality: 1723.50 Length: 455
Ratio: 4.113 Gaps: 1
Percent Similarity: 92.088 Percent Identity: 72.527

alignment_block:
US-09-589-510-4 x AAX18199 ..
Align seg 1/1 to: AAX18199 from: 1 to: 1730
1 MetArgIleGlnIlyValGlnSerThrSerLysLysGlnArgIleAlaIarh 17
|||||

```

```

60 ATGAAGATTGAGAGGTGAGAGCACTACGAAGACGACGCCATCGCCTC 109
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnIlyMetAlaIleA 34
|||||
110 CCACAGCCACAGTGAAGAGGCTGGGCTGGACGACGAGCGCTTGGCCAAAGC 159
34 lAlaValAlaIleGlyPheValGlyGlnAlaAlaAlaArgGlnAlaIleGly 50
|||||
160 AGCGCGCCTCAGCGCTGTGGGCCAGAGAACGCCGACGAGCATGTGCGC 209
51 LeuAlaValAspMetIleArgGlnIlySerMetAlaGlyArgAlaValle 67
|||||
210 GTTCATGTAGTAATTAATCAAAAGCAAGAAATGGCTGGAAGAGCTGTCTT 259
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
|||||
260 GTTGGCAGAGACTCTCTGCACTGGCAAGACAGCTCTGTGCTGTGCTATTTG 309
84 lAGlnGlnLeuGlySerLysValProPheCySPrometValGlySerGlu 100
|||||
310 CTCAGAGAGCTGGGTAGTAAGGTCCCTCTGCAATGTGGGAGGTGA 359
101 ValTYrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
|||||
360 GTTACTCACTGAGATCAGAGACAGAGGTGCTGATGAGAACTTCCG 409
117 gArgAlaIleGlyLeuArgIleLysGlnAsnLysGlnValTYrGlnIlyG 134
|||||
410 CAGGGCCATTTGGCTGCGAATTAAGAGACCAAGAGAACTTATAGAGGTG 459
134 lUValThrGlnLeuSerProGlnIlyAlaIlySerThrThrGlyIlyTYr 150
|||||
460 AAGTCACAGAGCTAACCTCGTGTAGACAGAGAAATCCATGGAGGATAT 509
151 AlAlySerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
|||||
510 GGCAAACCAATTAAGCATGTATGACTCAAAACAGCAAAAGAAC 559
167 rLysGlnLeuLysLeuAspSerSerIleTYrAspAlaLeuIleLysGln 184
|||||
560 CAACAGTTGAACACTGCAGCCCGACATTTTGAAAGTTTGCAAGAAAGAC 609
184 ySValAlaValGlyAspValIleTYrIleGlnAlaAsnSerGlyAlaVal 200
|||||
610 GAGTGAAGCTGTGAGATGTGATTACATTGAAGCAACAGTGGCGCTG 659
201 LysArgValGlyArgCyAspSerPheAlaThrGlyTYrAspLeuGlnAl 217
|||||
660 AAGAGGCGAGGCGAGGTGTGATCTATGTCACAGAAATTCGACTTGAAGC 709
217 aGlnIlyTYrValProIleProLysGlyGlnValHisLysLysGlnI 234
|||||
710 TGAAGATGTATGCTCCCTTGCCAAAAGGGATGTGCACAAAAGAAAGAA 759
234 lValAlaAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
|||||
760 TCATCCAGATGTGACTTGTGATGCTGTGATGTGCTTAATAGCCGCGCC 809
251 GlnIlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267
|||||
810 CAGGGGGGACAAAGATATCTGTGCATGTGGCCAGCTAAAGAGCCAAA 859
267 GlyThrGlnIleThrGlnLysLeuArgGlnGlnIleAsnLysValAla 284
|||||
860 GAAGACAAATATCACAGCAAACTTCCAGGGGAATTAATTAAGTGGTGA 909
284 snArgTYrIleAspGlnIlyIleAlaGlnLeuValProGlyValLeuPhe 300
|||||
910 ACAAGTACATGACACAGGCGCATGTGCTGAGTGTGCGGAGTGTCTGTT 959
301 lLeaspGlnValHisMetLeuAspIleGlnLysPheSerTYrLeuAsnAr 317
|||||
960 GTTGATGAGTTCACATCTGTGACATTAAGTGTTCACCTTACCTGCACCG 1009

```







```

|||||
332 CTCAGAGCTGGTGTAGTAGTCCCTCTGCGCAATGCTGGGAGGAA 381
101 VALTYRSErSerGIuValLysLysThrLValLeuMetGIuAsnPheAr 117
382 GTTTCCTCACTGAGATCAGAGAGAGAGGTGCTGATGGAGAACTTCCG 431
117 gATgAlaIleGIuLeuArGIleLysGIuAsnLysGIuValTYrGIuLg 134
432 CAGGCGCATTTGGCTGCGAATTAAGAGACCAGGAAGTTATGAGGTG 481
134 LuValThrGIuLeuSerProGIuGIuAlaGIuSerThThGIuGIuTYr 150
482 AAGTCACAGAGCTAACTCCGTGTGAGACAGAGAAATCCCATGGAGAT 531
151 AAlaLysSerLLeSerHisValIleLeuSerLeuLysThrValLysGI 167
532 GGCAGAAACCTTAGCCATGTGATCATAGCACTCAAAACAGCCAAAGAAC 581
167 rLysGIuLeuLysLeuAspSerSerLLeTYrAspAlaLeuIleLysGI 184
582 CAACAGCTTGAACACTGGACCCAGCATTTTGAAGTTTGCAGAAAGAGC 631
184 ySValAlaValGIuAspValIleTYrIleGIuAlaAsnSerGIuAlaVal 200
632 GAGTAGAAGCTGGAGATGATTTACATTGAGCCAAACAGCTGGGCGCTG 681
201 LysArGVAlGIuArGysAspSerPheAlaThrGIuTYrAspLeuGIuAl 217
682 AAGAGCGAGGCGAGTGTGATCTACTGCGACAGAAATTCGACTTGAAGC 731
217 aGIuGIuTYrValProIleProLysGIuGIuValHisLysLysGIu 234
732 TGAAGATATGTCCCTGCCAAAGGAGATGTCACAAAGAAAGAA 781
234 LeValGIuAspValThrLeuHisAspLeuAspAlaAlaAsnLagInPro 250
782 TCATCCAGATGTGACCTTGCATGACTGTGATGTGCTAATCGCGGCC 831
251 GIuGIuLysGIuAspIleLeuSerLeuMetGIuGIuMetLysProAr 267
832 CAGGGGGCAGAGATATCCGTGCATGATGGCGACCTAATCAAGCCAA 881
267 gLysThrGIuIleThrGIuLysLeuArGIuGIuIleAsnLysValAla 284
882 GAAGACAGAAATCAGACAACTCGAGGAGATTAAAGTGTGA 931
284 snArGIuTYrLeuAspGIuGIuIleAlaGIuLeuValProGIuValLeuPhe 300
932 ACAAGTACATGACCGAGGCAATGCTGAGCTGCTCCGGGTGCTGTT 981
301 IleAspGIuValHisMetLeuAspIleGIuLysPheSerTYrLeuAsnAr 317
982 GTTGATGAGTCCACATGCTGAGCATGTGAGTCTCACTACCTGCACCG 1031
317 gAlaLeuGIuSerProLeuSerProIleValIleLeuAlaThrAsnArG 334
1032 CGCCCTGAGCTTCTATCGCTCCATCGTCACTTGTTCATCCACCGAG 1081
334 LyIleCysAsnValArGIuTYr...AspMetThrSerProHisGIuLe 349
1082 GCAACTGTGCTCATGAGGAGCAGTGAACATCAATCCCTTCACGGCAGTC 1131
350 ProValAspLeuAspArGIuValIleLeuArGIuThrGIuTYrGI 366
1132 CCTTTCAGCTTCTTGACCGAGTATATATCCGACCTGCTGTATAC 1181
366 yProThrGIuMetIleGIuIleLeuAlaIleArGIuAlaGIuGIu 383
1182 TCACAGAGAAATGAACAGATCATTAATAATCCGCGCCACAGAGAA 1231
383 LeAspMetAspGIuGIuSerLeuAlaTYrLeuGIuGIuIleGIuGIu 399
|||||

```

```

1232 TCACATCATGATGAGGAGGCACTGAACACACTGGGGAGATTGGACCCAG 1281
400 ThrSerLeuArGIuHisAlaIleGIuLeuIleSerProAlaSerValSe 416
1282 ACCACACTGAGTACTCAGTCACTGCTGACCCCGGCAACTTGTCTGC 1331
416 rLysThrAsnGIuArGIuLysIleCysLysAlaAspLeuGIuValS 433
1332 TAAATCAACGGGAAAGCAGCATTTGACAAAGAGCATGTCAACAGATCA 1381
433 erGIuLeuTYrLeuAspAlaLysSerSerAlaArGIuLeuGIuGIuIn 449
1382 GTGACTTTTCTATGATGCCAAGTCTCCGCCAAATCTGCTGACACG 1431
450 GIuGIuArGIuTYrIle 454
1432 CAGGATTAAGTACATG 1446

seq_name: /SIDSI/gcgdata/geneseq/geneseq/NN2000.DAT:AA059220
seq_documentation_block:
ID AAC59220 standard; CDNA: 1744 BP.
XX
XX AAC59220;
XX
XX 29-JAN-2001 (first entry)
XX
DE Human secreted protein cDNA sequence #6.
XX
XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; ss.
XX
XX Homo sapiens.
XX
XX WO20005199-A1.
XX
XX 21-SEP-2000.
XX
XX 09-MAR-2000; 2000MO-US06014.
XX
XX 12-MAR-1999; 99US-0124095.
XX 11-JUN-1999; 99US-0138598.
XX 03-DEC-1999; 99US-0168665.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI: 2000-572359/53.
XX P-PSDB; AAB27799.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX PT used in preventing, treating or ameliorating a medical condition
XX
XX Claim 1; Page 343; 433pp; English.
XX
XX The invention relate to the isolation of genes AAC59215-C59261 encoding
XX CC 47 human secreted proteins AAB27794-B27840. The genes can be used to
XX CC generate fusion proteins by linking to the gene for the human
XX CC immunoglobulin G Fc portion for increasing the stability of
XX CC the fusion protein as compared to the human protein only. The genes and
XX CC proteins are useful for preventing, ameliorating or treating medical
XX CC conditions, e.g. by protein or gene therapy. The genes are isolated
XX CC from a range of human tissues disclosed in the specification. The
XX CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
XX CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

```

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias; (d)  
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.

XX Sequence 1744 BP; 476 A; 403 C; 480 G; 384 T; 1 other;

# alignment\_scores:

Quality: 1570.00 Length: 458  
 Ratio: 3.811 Gaps: 3  
 Percent Similarity: 89.956 Percent Identity: 70.306

# alignment\_block:

US-09-589-510-4 x AAC59220 ..

Align seg 1/1 to: AAC59220 from: 1 to: 1744

```

1 MetArgIleGluGluValGlnSerThrSerLysGlnArgIleAlaThr 17
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 ATGAGATTGAGGAGTGAAGAGCACTACGACGCGCATCGCTC 125
17 rHstHrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIle 34
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
126 CCGAGCGCAGGTGAAGAGCGCTGGCGACGAGCGGCTGGCCAAGC 175
34 LalaValAlaGlyPheValGlyGlnAlaAlaArgGlnAlaGly 50
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
176 AGCGGGCTCAGGGCTTGTGGCGCAGAGACGCGGAGAGCATGTGGC 225
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaVal 67
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
226 GTCATGATTGAATTAATCAAAAGCAAAATGCTGGAAGAGCTGTCTT 275
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaGlyIle 84
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
276 GTTGGCAGAGACCTCCTGGAACCTGGCAAGACAGCTGTGGCTATTG 325
84 IalGlnIleuLysSerLysValProPheCysProMetValGlySerGlu 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
326 CTCAGAGAGTGGGTAGTAGTCCCTCTGCCAATGGTGGGAGTGAA 375
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPhe 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
376 GTTACTCTACACTGATCAAGAAAGACAGAGCTGATGGAAGACTTCG 425
117 gArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
426 CAGGCC.ATTGGGCTGCGAATAAAGAGACCAAGAGATTATGAGT. 473
134 IuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyLys 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
474 AAGTCAACAGAGCTAATCTCGTGTGAGACAGAGAAATCCCATGSGAGATAT 523
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGly 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
524 GCGAAACCATTAAGCCATGTGATCATAGACTCAAAACAGCCAAAGGAG 573
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
574 CAAACAGTTGAAACTGAGCCCGCATTTTGAAGTTTGCAGAAAGAGC 623
184 ysValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAla 200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
624 GAGTGAAGAGCTGGAGATGTGATTACATTGAAGCCAAACAGTGGCGCTG 673
201 LysArgValAlaGlyArgCysAspSerPheAlaThrGluLysIleAsnGlu 217
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
674 AAGAGCGAGGCGAGGTGATCACTATGCGCACAGAAATTCGACTTGAG 723
217 aGluGluTyrValProIleProLysGlyGluValHisLysLysGluI 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

724 TGAAGATATGTCCTTGGCCAAAAGGGATGTGCACAAAAGAAAGAA 773
234 IeValGlnSerValThrLeuHisAspLeuAspAlaAlaAsnIlePro 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
774 TCATCCAGATGTGACCTTGACATGACTTGATGTGCTATATGCGGGCC 823
251 GlnGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
824 AG. GGGGAGCAAGATATCCTGTCATGATGGGCGACGTAAAGACCAAG 872
267 GlyThrGluIleThrGluLysLeuArgGln.....GluLeuAsnL 281
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
873 AAGA.....CAGAAATCAGACAAACTTCAGGGAGATTATA 910
281 ysValValAsnArgTyrIleAspGluGlyIleAlaGluLeuValProGly 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
911 AGGTGTGAAACAAGTACATCGACAGGCGATGTGAGCTGTGCTCCGGGT 960
298 ValLeuPheIleAspGluValHisMetLeuAspIleGluCysPheSer 314
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
961 GTGCTGTTGTTGATGAGGTCCATGCTGGACATGAGTGCCTCACCTA 1010
314 rLeuAsnArgAlaLeuGluSerProLeuSerProIleValIleLeuAla 331
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1011 CCGACACCGCGCCCTGGAGTCTTCTATGCTGCCATCTCATCTTTGCAT 1060
331 hrAsnArgGlyIleCysAsnValArgGlyThr...AspMetThrSerPro 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1061 CCAACCGAGGCAACTGTGTCATCAAGGCACTGAGAGATCATCATCCCT 1110
347 HisGlyIleProValAspLeuAspArgLeuValIleIleArgThrGlu 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1111 CAGCGCATCCCTCTTGRCTTCTGACCGAGATGATTAATCCGGACCAT 1160
363 uThrTyrGlyProThrGluMetIleGlnIleLeuAlaIleArgAlaGln 380
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1161 GCTGATATCTCCACAGAAATGAATCATTAATAATCGTCCCA 1210
380 aGluGluIleAspMetAspGluGluSerLeuAlaTyrLeuGlyGluIle 396
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1211 CGGAAGC.ATCAACATCAAGTAGAGAGGCACTGAACCACTGGGGAGATT 1259
397 GlyGlnGlnThrSerLeuArgHisAlaIleGlnLeuLeuSerProAla 413
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1260 GGCACCAAGACACACACTGAGGTACTCAGTGAAGTGTGACCCCGCAA 1309
413 rValValSerLysThrAsnGlyArgGluLysIleCysLysAlaAspLeu 430
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1310 CTTCCTTGCTAAATCAACGGGAGAGCACAGCATTTGAAGAGCATGTCG 1359
430 IuGluValSerGlyLeuTyrLeuAspAlaLysSerSerAlaArgLeu 446
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1360 AAGGATCAAGAGACTTTTCTATGATGCTCAAGTCTCCGCAAAATCTCG 1409
447 GlnGluGlnGlnGluArgTyrIle 454
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1410 GCTGACCAAGCAGATAGTACATG 1433

```

seq\_name: /stds1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ52915

seq\_documentation\_block:

ID AAZ52915 standard; cDNA; 1509 BP.

XX AAZ52915;

XX 14-MAR-2000 (first entry)

XX Human prostate tumor cDNA library derived EST fragment #58.

XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;

XX treatment; ds.

XX Homo sapiens.

OS

```
XX DEL9820190-A1.
PN
XX 04-NOV-1999.
PD
XX 28-APR-1998; 98DE-1020190.
PF
XX 28-APR-1998; 98DE-1020190.
PR
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
PA
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
PI WPI: 1999-621386/54.
XX P-PSDB: AAY73986, AAY73987, AAY73988.
DR
XX New human nucleic acid sequences from pancreatic tumors, and related
PT proteins
PS Claim 2; Page 229-230; 502pp; German.
XX
CC This invention describes novel polypeptides and their encoding nucleic
CC acids derived from human pancreatic tumor tissue which have cytostatic
CC activity. The sequences are also useful in producing pharmaceutical
CC compositions for treatment of pancreatic tumors. AA52858-253014
CC represent expressed sequence tag (EST) fragments derived from a human
CC pancreatic tumor cDNA library and which encode the proteins represented
CC in AAY73814-74252.
XX
SQ Sequence 1509 BP; 416 A; 338 C; 416 G; 339 T; 0 other;

alignment_scores:
    Quality: 1507.50      Length: 401
    Ratio: 4.074          Gaps: 1
Percent Similarity: 92.269 Percent Identity: 73.566

alignment_block:
US-09-589-510-4 x AA52915 ..

Align seq 1/1 to: AA52915 from: 1 to: 1509

47 GUAAlaLaGlyLeuAlaValaAspMetIleArgInLysLysMetAlaG1 63
   ::::::::::::::::::::|::::::::::::::::::|
53 CAGGCACTGGCGCTCATGATGATTAATCAAAAGCAAAATGGCTGCG 102
   |::::::::::::::::::|
63 YARGAlaValLeuLeuAlaGlyProProAlaThrGlyLysThrAlaLeuA 80
   |::::::::::::::::::|
103 AGGAGCTGCTGTGTGGCAGGACCTCTGGAAGCTGCAAGACAGCTCTGG 152
   |::::::::::::::::::|
80 LaleuGlyIleAlaGlnGlnLeuGlySerLysValProPheCysPromet 96
   |::::::::::::::::::|
153 CTCTGGCTATTGCTCAGAGCTGGGTAGTAAAGTCCCTTCTGCCCAATG 202
   |::::::::::::::::::|
97 ValGlySerGluValLysSerSerGluValLysLysThrGluValLeuMe 113
   |::::::::::::::::::|
203 GTGGGAGCTGAAGTTTACTCACTGAGATCAAGAAGCAGAGTCTGATN 252
   |::::::::::::::::::|
113 LgluAsnPheArgArgAlaIleGlyLeuArgIleLysGluAsnLysGluV 130
   |::::::::::::::::::|
253 GGAGAACTTCGCGAGGCCATTTGGGCTGCGAATTAAGGAGAACCAAGAG 302
   |::::::::::::::::::|
130 aLTYrGlnGluGluValThrGlnLeuSerProGlnGluAlaGlnSerThr 146
   |::::::::::::::::::|
303 TTTTGAAGGTGAAGTCAACAGCTAACTCCGTGTGAGACAGAAATCCG 352
   |::::::::::::::::::|
147 ThrGlnGlyTyrAlaLysSerLysSerHisValIleLysSerLeuLysTh 163
   |::::::::::::::::::|
353 ATGGAGAGATATGGCAAAACCATTAAGCATGTGATCATAGACTCAAAAC 402
   |::::::::::::::::::|
163 rValLysGlyThrLysGlnLeuLysLeuAspSerSerLysThrAspAlaL 180
   |::::::::::::::::::|
403 AGCCAAAGGAACCAACAGTTGAACGTGAGCCCAAGCATTTTGAAGATT 452
```

```
180 euLleLysGlnLysValAlaValaGlyAspValIleTyrIleGlnAlaAsn 196
   ||::::::::::::::::::|
453 TGCAGAAAGAGCGAGTAGAAGAGCTGAGATGTGATTTACATTTGAAGCCAAC 502
   ||::::::::::::::::::|
197 SerGlyAlaValLysArgValaGlyArgCysAspSerPheAlaThrGlyTyr 213
   ||::::::::::::::::::|
503 ACTGGGGCCGTGAAGAGCCAGCGAGGTGTGATCTATGCTCCACAGAAAT 552
   ||::::::::::::::::::|
213 rAspLeuGlnAlaGlnGluTyrValProIleProLysGlyGluValHisL 230
   ||::::::::::::::::::|
553 CGACCTTGAGCTGAAAGAGTGTGCTCCCTTG.CCAAAAGGGAGATGTCACA 601
   ||::::::::::::::::::|
230 yLysLysGlnLleValaGlnAspValThrLeuHisAspLysAlaAla 246
   ||::::::::::::::::::|
602 AAAAGAAAGAAATCATCCACATGTGACCTTCATGACATTTGGATGTGGCT 651
   ||::::::::::::::::::|
247 AsnAlaGlnProGlnGlyGlnAspIleLeuSerLeuMetGlyGlnMe 263
   ||::::::::::::::::::|
652 AATGGGGGGCCCCAGGGGGAACAAAGATATCCTGTTCATGATGGCGCAGCT 701
   ||::::::::::::::::::|
263 tMetLysProAlaGlyThrGlnIleThrGlnLysLeuArgGlnGlnIleA 280
   ||::::::::::::::::::|
702 AATGAGCCCAAGAGACAGAAATCAACAGCAAACTTGAGAGGAGATTA 751
   ||::::::::::::::::::|
280 snLysValValAsnArgTyrIleAspGlnGlyIleAlaGlnLeuValPro 296
   ||::::::::::::::::::|
752 rTRAGGTGTGAACAGATACATGACGACAGGGGATTTGCTGACCTGCTCCG 801
   ||::::::::::::::::::|
297 GlyValLeuPheIleAspGluValHisMetLeuAspIleGlyCysPhe 313
   ||::::::::::::::::::|
802 GGTGTCTGTTGTGTGATGAGGTGCACATGCTGGACATTTGATGCTTTCAC 851
   ||::::::::::::::::::|
313 rTyrLeuAsnArgAlaLeuGlnSerProLeuSerProIleValIleLeuA 330
   ||::::::::::::::::::|
852 CTACCTGACACCGCGCCCTTGAGCTTTATGCTGCCATTCGTCATCTTGG 901
   ||::::::::::::::::::|
330 lATHrAsnArgGlyIleCysAsnValArgGlyThr...AspMetThrSer 345
   ||::::::::::::::::::|
902 CATCCAAACGAGCACTGTGTATCAGAGGCACTGAGGACATCATCCATCC 951
   ||::::::::::::::::::|
346 ProHisGlyIleProValAspLeuLeuAspArgLeuValIleIleArgTh 362
   ||::::::::::::::::::|
952 CTTCAGGGCATCCCTCTTGACCTTCTGGACCCAGTATGATATATCCGGAC 1001
   ||::::::::::::::::::|
362 rGluThrTyrGlyProThrGlnMetIleGlnIleLeuAlaIleArgAlaG 379
   ||::::::::::::::::::|
1002 CATGCTGTATATCCACAGGAATGAACAGATCATTAATAATCCGTGCC 1051
   ||::::::::::::::::::|
379 lValaGlnGluIleAspMetAspGlnGlnSerLeuAlaThrLeuGlnGlu 395
   ||::::::::::::::::::|
1052 ACAGCGAAGGAATCAACATCAGTAGAGAGGCACTGAAACCACTGGGGGAG 1101
   ||::::::::::::::::::|
396 lIleGlyGlnGlnThrSerLeuArgHisAlaIleGlnLeuIleSerProAl 412
   ||::::::::::::::::::|
1102 ATTGGACCAAGACACACATGAGGTACTGATGCTGACCTGATCCCGGCG 1151
   ||::::::::::::::::::|
412 aSerValValSerLysThrAsnGlyArgGlnLysIleCysLysAlaAspL 429
   ||::::::::::::::::::|
1152 CAACCTGCTTGTCTTAATAATCAACGGGAAGCACACCATTTGAGAAAGACATG 1201
   ||::::::::::::::::::|
429 euGlnGluValSerGlyLeuThrLeuAspAlaLysSerSerAlaArgLeu 445
   ||::::::::::::::::::|
1202 TCGAAGAGATCAGTGAACCTTTCTATGATGTCACAAAGTCTCTCGCCAAATTC 1251
   ||::::::::::::::::::|
446 Leu 446
   ||
1252 CTG 1254

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF13700
seq_documentation_block:
ID AAF13700 standard; cDNA; 1213 BP.
```

XX AAF13700;  
 AC 13-MAR-2001 (first entry)  
 DT Aspergillus oryzae EST SEQ ID NO:6223.  
 DE  
 XX Multiple gene expression; filamentous fungal cell; EST;  
 XX expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
 KW culture condition; environmental stress; spore morphogenesis;  
 KW metabolic pathway engineering; catabolic pathway engineering; ss.  
 XX  
 OS Aspergillus oryzae.  
 XX  
 PN WO200056762-A2.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 22-MAR-2000; 2000WO-US07781.  
 XX  
 PR 22-MAR-1999; 99US-0273623.  
 XX  
 PA (NOVO ) NOVO NORDISK BIOTECH INC.  
 PA (NOVO ) NOVO NORDISK AS.  
 XX  
 PI Berka RM, Rey WM, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;  
 DR WPI; 2000-594572/56.  
 XX  
 PT Monitoring differential expression of genes in filamentous fungal cells  
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
 PT substrate of expressed sequence tags -  
 XX  
 PS Claim 88; Page 2561; 3161pp; English.  
 XX  
 CC The present invention describes a method for monitoring differential  
 CC expression of genes in a first filamentous fungal (FF) cell relative to  
 CC expression of the same genes in one or more second filamentous fungal  
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
 CC are used in the methods for monitoring differential expression of genes  
 CC in a first filamentous fungal (FF) cell relative to expression of the  
 CC same genes in one or more second filamentous fungal cells. Monitoring  
 CC the global expression of genes from FF cells allows the production  
 CC potential of the microorganisms to be improved. New genes may be  
 CC discovered, possible functions of unknown open reading frames can be  
 CC identified and gene copy number variation and stability can be  
 CC monitored. The expression of genes can be used to study how FF cells  
 CC adapt to changes in culture conditions, environmental stress, spore  
 CC morphogenesis, recombination, metabolic or catabolic pathway  
 CC engineering. Using ESTs provides several advantages over genomic or  
 CC random cDNA clones including elimination of redundancy as one spot on an  
 CC array equals one gene or open reading frame, and organization of the  
 CC microarrays based on function of the gene products to facilitate  
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from  
 CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from  
 CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus  
 CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are  
 CC all specifically claimed in the present invention.  
 XX  
 SQ Sequence 1213 BP; 336 A; 325 C; 275 G; 276 T; 1 other;

## alignment\_scores:

Quality: 1064.50 Length: 297  
 Ratio: 3.972 Gaps: 1  
 Percent Similarity: 90.236 Percent Identity: 68.687

## alignment\_block:

us-09-589-510-4 x AAF13700 ..  
 Align seg 1/1 to: AAF13700 from: 1 to: 1213

117 ArgArGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGlu 133  
 1 CGAGAGCGCATGTGGTCTCCGTCGCGTAACCAAGAAAGAGTACCGAAG 50  
 133 YGluValThrGluLeuSerProGluGluAlaGluSerThrThrGlyGlyT 150  
 51 CGAAGTCACGACCTTAACGCTGAGCAAACTGAAATCATTTGGAGGCT 100  
 150 YrAlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGly 166  
 101 ATGACCCCAATCATGACCTGATTTATGATGATGATGATGATGATGAT 150  
 167 ThrLysGluLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 183  
 151 ACCAAGACGCTGCGCTGATCCAGCATTTTGAAGGCTATTCANAAAGA 200  
 183 ULysValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaV 200  
 201 ACAGTCTACTGTGGAGATGCTATCTACATCGAACGAACTGGTGCTT 250  
 200 aLysArGValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGlu 216  
 251 GTAAGAGAGTGGCCGATGCCATGCCATGCGACCGAGTTCGACCTTAA 300  
 217 AlaGluGluTyrValProIleProLysGluGluValHisLysLysGlu 233  
 301 GCAGAAAGTACGTTCTGTGTAACCAAGGAGGTTTCACAGAAAGAAAGA 350  
 233 uLleValGlnAspValThrLeuHisAspLeuAspAlaLysAsnLysGln 250  
 351 AATCGTACAGGACGTGACGCTACATGACCTGACATGCGACCTCGGC 400  
 250 rGlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysPro 266  
 401 CACAAGTGGACAGGACGCTATGACATGATGAGCAACATGATGAACCC 450  
 267 ArgLysThrGluIleThrGluLysLeuArgGlnLysLysValVal 283  
 451 AAGAAGACCGAGATCGGATGACCTGATGAGATGAGATGAGATGAG 500  
 283 lAsnArgTyrIleAspGluGluGlyIleAlaGluLeuValProGluValLeu 300  
 501 TAGCGGTATATCGAACAGAGATGCTGAACTGGTCTCTGCTCTCT 550  
 300 heLleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsn 316  
 551 WTATTGATGAGGTCCACATGTTGGATGATGATGATGATGATGATGAT 600  
 317 ArgAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsn 333  
 601 CGGCGACTGTGAATCTTCATCTCCCATTTGATGCTGCTCCACCG 650  
 333 gGlyIleCysAsnValArgGlyThr...AspMetThrSerProHisGlyT 349  
 651 CGGCGACACTGTGATCGCGGACCGGACGACATCAAGCGGACGAGGA 700  
 349 leProValAspLeuLeuAspArgLeuValIleLeuArgThrGluThrTyr 365  
 701 TTCCCTGTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750  
 366 gLyProThrGluMetIleGlnIleLeuAlaIleArgAlaGluValGluGlu 382  
 751 TCTTCGAGAGATCAAGACCATCTTCGCTTCGCGCCAAAGACAGAGAG 800  
 382 uLleAspMetAspGluGluSerLeuAlaTyrLeuGluGluIleGlyGln 399  
 801 TCTCAACATTCACGATCCGCTCGACAAAGATTCGCGACGACGACGA 850  
 399 lThrSerLeuArgHisAlaIleGlnIleLeuIleSerProAla 412  
 851 ACGTCAAGCTTGGCGTACGACATGACATTTGACCCGACGCT 891

```

seq_name: /SIPSI/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90316
seq_documentation block:
ID AAA90316 standard; cDNA; 1492 BP.
AC AAA90316;
XX
XX 22-DEC-2000 (first entry)
XX
XX Human HEL50 DNA helicase cDNA.
XX
XX Human; HEL50; DNA helicase; TIP49 homologue; EST AA374580;
XX expressed sequence tag; antibody; antisense therapy; diagnosis;
XX helicase-associated disease; genetic disorder; tumour; cancer; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 1..1392
XX FT CDS
XX FT /*tag= a
XX FT /product= "Human HEL50"
XX FT /function= "DNA helicase"
XX
XX MO200047731-A1.
XX
XX 17-AUG-2000.
XX
XX 22-NOV-1999; 99WO-JP06519.
XX
XX 10-FEB-1999; 99JP-0033062.
XX
XX (SUME ) SUMITOMO ELECTRIC IND CO.
XX
XX Tamura T;
XX
XX WPI: 2000-543587/49.
XX
XX DR P-PSDB; AAB21093.
XX
XX DNA helicase of human origin, antibodies to it and DNA encoding it for
XX investigation diagnosis and treatment of genetic disorders involving
XX helicase gene expression
XX
XX Claim 6; Page 41; 49pp; Japanese.
XX
XX This sequence represents cDNA encoding a novel human DNA helicase,
XX HEL50, which has homology with a part of human TIP49. cDNA encoding
XX human HEL50 was initially identified in an EST (expressed sequence tag)
XX database as EST AA374580. The invention relates to human and yeast HEL50
XX cDNA helicases (AAB21093, AAB21094), and to cDNAs encoding them
XX (AAA90316, AAA90317). It also encompasses HEL50 antisense
XX oligonucleotides and anti-HEL50 antibodies. HEL50 nucleotides, proteins
XX and antibodies may be used for the investigation, diagnosis and treatment
XX of diseases with which DNA helicase is associated, such as genetic
XX disorders involving variant forms of the gene, and tumours in which
XX altered levels of expression of the DNA helicase gene occur.
XX
XX Sequence 1492 BP: 352 A; 439 C; 454 G; 247 T; 0 other;

alignment_scores:
Quality: 973.00 Length: 455
Ratio: 2.845 Gaps: 4
Percent Similarity: 75.165 Percent Identity: 43.956

alignment_block:
US-09-589-510-4 x AAA90316 ..
Align seq 1/1 to: AAA90316 from: 1 to: 1492
2 ArgIleGluGluValGlnSerThrSerLysGlnArgIleAlaThrH 18
: : : : : | | | | | : : : : : : : : : : : : : : : :
25 AAAGTCCCGGAGATCCGTGATGTACACAGATTGAGCGAATCGTGCCCA 74

```

```

18 sThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleAla 35
: : : : : | | | | | : : : : : : : : : : : : : : : :
75 CTCCACATCTCCGGGAGCTGGGCTGAGCATGCTTGGAGCCTCGGCAGG 124
35 euAlaIaGlyPheValGlyGlnAlaAlaIaArgLysAlaIaGlyLeu 51
: : : : : | | | | | : : : : : : : : : : : : : : : :
125 CTTCGCAAGGAGCATGTGGTGTACCTGGCGGCACGCGGCGGCTGGCGT 174
52 AlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaValLeu 68
: : : : : | | | | | : : : : : : : : : : : : : : : :
175 GTGCTGGAGATGATCCGCAAGGAGATGTCGCGTGGCGAGCTCTTAT 224
68 uAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleAla 85
: : : : : | | | | | : : : : : : : : : : : : : : : :
225 TGTGTGGCCAGCCGGGCACGGGGAAGACGGCATTCGCTGATGGCATGG 274
85 LngLysGlySerLysValProPheCysProMetValGlySerGluVal 101
| | | | | : : : : : : : : : : : : : : : :
275 AGGCCCTGGGCCCCGACACCGCATTCACAGCCATCCGCGGAGTGAATC 324
102 TySerSerGluValLysLysThrGluValLeuMetGluAsnPheArg 118
: : : : : | | | | | : : : : : : : : : : : : : : : :
325 TTCTCCTGGAGATGAGCAAGACCGAGCGCTGACGAGCCTTCCGCGG 374
118 gAlaIleGlyLeuArgGlyLeuGlyGluAsnLysGluValTyrgLysGlu 135
: : : : : | | | | | : : : : : : : : : : : : : : : :
375 GTCCATCGGCGCTTCGATCAAGAGAGACGAGATCATGCAAGGGAGG 424
135 AlThrGluLeuSerProGluGluValGlnSerThrThrGlyGlyTyra 151
: : : : : | | | | | : : : : : : : : : : : : : : : :
425 TGTGTGAGATCCAGATTGATGACACGACGAGGAGGCGGTCCAAAGTG 474
152 LysSerIleSerHisValIleIleSerLeuLysThrValLysGlyThr 168
: : : : : | | | | | : : : : : : : : : : : : : : : :
475 GGCMAA.....CTGACCTCAAGACACAGCAGATGAGGAGAC 509
168 sGlnLeuLysLeuAspSerSerIleTyraSpAlaLeuIleLysGlyLys 185
510 CATCTACGACCTGGGACCAAGATGATTGATCCCTCCAGCAAGCAAG 559
185 AlAlaValGlyAspValIleTyrlleGluAlaAsnSerGlyAlaVal 201
| | | | | : : : : : : : : : : : : : : : :
560 TCCAGCGCGGGAGCGTGTATCACCATTCACAGCGGAGCGCAAGATCTCC 609
202 ArgValGlyArgCysAspSerPheAlaThrGluTyraSpLeuGluAla 218
: : : : : | | | | | : : : : : : : : : : : : : : : :
610 AAGCTGGGCGCTCTCTTCACACGCGCCGCGACTACAGCTATGGGCTC 659
218 uGlu.....TyraLProIleProLysGlyGluValHisLysLys 233
: : : : : | | | | | : : : : : : : : : : : : : : : :
660 CCAGACCAAGTTCTGTCAGTCGCCAGATGGGAGAGCTCCAGAAAGCAAG 709
233 LuIleValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGln 249
: : : : : | | | | | : : : : : : : : : : : : : : : :
710 AGTGTGTACACCGCTGTCCCTGCACGAGATGACGATCATCACTCTGCG 759
250 ProGlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLys 266
| | | | | : : : : : : : : : : : : : : : :
760 ACCCAGGCG.....TTCCTGGGCGCTC.....TTCCTCAGG 788
266 cArgLysThrGluIleThrGlnLysLeuArgGlnGluLeuLysVal 283
789 TGACACAGGAGATCAAGTCAAGTCCGTGAGGAGATCAATGTCGAAG 838
283 AlAsnArgTyrlleAspGlnGlyIleAlaGluLeuValProGlyValLeu 299
| | | | | : : : : : | | | | | : : : : : : : : : : :
839 TGGCTGATGTGGCGGAGGAGGAGGAGGAGATCATCCTGAGTGTCTG 888
300 PheIleAspGlnValHisMetLeuAspIleGlyCysPheSerTyrlleAs 316
: : : : : | | | | | : : : : : : : : : : : : : : : :
889 TTCATCTGACGAGGTCCACATGCTGTGAGATCGAGACTTCTCTCTCAA 938
316 nArgAlaLeuGlnSerProLeuSerProIleValIleLeuAlaThrAsn 333

```

```

|||||.....|
939 CCGGGCCCTGAGATGACATGGCGCTGTCGATCATGCGCACCAACC 988
333 rgglyilecysasnvalarglythrasmethrserprohisglyile 349
989 GTGGCATCAGCCGAATCCGGGGACACGCTACCAAGAGCCCTCACGGCATC 1038
350 ProValAspleuLeuAsparGluValIleIleArgThrGluThrTyrGI 366
1039 CCATAGACCTGCTGACCGCGTGTATGCTCTCCACCCACCCCTCAAG 1088
366 yproThrGluMetIleGlnIleLeuAlaIleArgIleGlnValGluGI 383
1089 CGAGAAAGACAGACGAGATCTCCGCAATCCGTCGAGAGAGAGATG 1138
383 leaSpMetAspGluGluSerLeuAlaTyrLeuGluIleGlyGlnGI 399
1139 TGGAGATGATGAGAGACGCTACACGCTGTGACCCGCTACGCTGAG 1188
400 ThrSerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValAla 416
1189 ACCTCACTGCGCTACCGCATCCAGCTCATCACAGCTGCGCTGGTG 1238
416 rLysThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluGluValS 433
1239 CCGGAAACGACAGGCTACAGAGTGCAGGTGATGACATCAAGCGGCT 1288
433 erGlyLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlu 449
1289 ACTCACTCTTCCTGAGACAGAGTCCCGCTCCACCGCATGAGAGATAC 1338
450 GlnGluArgTyrIle 454
1339 CAGGACGCTCTCTC 1353
seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAA90317

```

seq\_documentation\_block:

ID AAA90317 standard; cDNA; 1416 BP.

AAA90317;

22-DEC-2000 (first entry)

DE Yeast HEL50 DNA helicase cDNA.

XX Yeast; HEL50; DNA helicase; TTP49 homologue; antibody; antisense therapy;  
 KW diagnosis; helicase-associated disease; genetic disorder; tumour; cancer;  
 ss.

XX Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT CDS

FT 1..1416

FT /\*Lag= a

FT /product= "Yeast HEL50"

FT /function= "DNA helicase"

XX W0200047731-A1.

XX 17-AUG-2000.

XX 22-NOV-1999; 99WO-JP06519.

XX 10-FEB-1999; 99JP-0033062.

XX (SDME ) SUMITOMO ELECTRIC IND CO.

XX Tamura T;

XX WPI, 2000-543587/49.

XX P-PSDB; AAB21094.

PT DNA helicase of human origin, antibodies to it and DNA encoding it for  
 PT investigation diagnosis and treatment of genetic disorders involving  
 PT helicase gene expression

PS Claim 8; Page 45-46; 49pp; Japanese.

XX This sequence represents cDNA encoding a novel yeast DNA helicase,  
 CC HEL50, which has homology with a part of human TTP49. The invention  
 CC relates to human and yeast HEL50 DNA helicases (AAB21093, AAB21094), and  
 CC to cDNAs encoding them (AAA90316, AAA90317). It also encompasses HEL50  
 CC antisense oligonucleotides and anti-HEL50 antibodies. HEL50 nucleotides,  
 CC proteins and antibodies may be used for the investigation, diagnosis and  
 CC treatment of diseases with which DNA helicase is associated, such as  
 CC genetic disorders involving variant forms of the gene, and tumours in  
 CC which altered levels of expression of the DNA helicase gene occur.

XX Sequence 1416 BP; 466 A; 238 C; 320 G; 392 T; 0 other;

alignment\_scores:

Quality: 932.50 Length: 444

Ratio: 2.809 Gaps: 6

Percent Similarity: 74.775 Percent Identity: 44.595

alignment\_block:

US-09-589-510-4 x AAA90317

Align seg 1/1 to: AAA90317 from: 1 to: 1416

```

15 llaalathrhistrhstlelysglyleuGlyLeuAspAlaAsnGlyme 31
|||||.....|
58 ATTGCTGCCACATCCCATATACAGGTTAGTCTAGATGAACATGCA 107
31 lAlaIleAlaLeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluA 48
108 ACAAGTCCACATCCGAGGATGAGTGGGCAATTGCCAAGCCCGCTG 157
48 lAlaGlyLeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArg 64
158 CTGCTGCTGATATGTGAATGTGTAACAAATGCTCAACATGACATGAG 207
65 AlAValLeuLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAla 81
|||||.....|
208 GCTGTTTGGTAGCGGGCCCCCTCACTGTAAGACCCCTCTGCCAT 257
81 uGlyIleAlaGlnGluLeuGlySerLysValProPheCysPrometValG 98
258 GGGTGTTCACAGTCTGGGTAAAGATGATACCATTCATCTGTAATGCG 307
98 lYserGluValTyrSerSerGluValLysLysThrGluValLeuMetGlu 114
308 GCTCAGAAATCTTTCTTAGAATTGAGTAACTGAGCATACTCAAC 357
115 AsnPheArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyr 131
|||||.....|
358 GCTTTTAGAATCCATCGGTATCAAAATCAAGAGACACACAAATGAT 407
131 rGluGlyGluValThrGluLeuSerProGluGluAlaGluSerThrThg 148
408 TGAAGTGAAGCTGTGAATTCATTCAT.....AGATCTATTACG 451
148 lYgLyTyrAlaLysSerIleSerHisValIleIleSerLeuSerThrVal 164
|||||.....|
452 GTGACACAAACAAGA.....AAATTCACATTTAAACTACC 489
165 lYsgLyThrLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuI 181
490 GATATGGAACAATTTATGATTAAGCAATAAATGATGATGCGCTAAC 539
181 elYsgLysValAlaValGlyAspValIleTyrIleGluAlaAsnSerG 198
|||||.....|
540 TAAAGAAAGGTATGCTGCGCATGTATTCTATTGATTAAGCTAGTG 589

```



```

198  lynaavlllysarivaiglyarvcysaspserphelathrcluyrasp 214
590  ggaagattaccagaacttagcaga.....tcccttgctagatctnogat 633
215  leuqlu1a.....gluqlutryvalproleprolysclgyl 227
634  tatgatgccatggcgctgataccagatttggctcaatgcggaagcga 683
227  uvalhislyslslysgluilevalglaspvalthrleuhsaspleua 244
684  actccaaaaaagaaaacagtgcttcacacggtgcactgcatagaattg 733
244  spalaalasnalaglinproclnglylgylgnasplleuserleumet 260
734  atgttatattatccagaacacagaa.....tttttggcatra... 771
261  glylmetmetlyspthrqlgylthrqluethrqluylsleuarqgl 277
772  .....tttacggtgacaccggtagaatttagctagagtagaaga 812
277  ngulileasnlsvaivaasnarytrrleaspluglyilealaglul 294
813  ccagattaaacacaaagtcgacgatggaaagaaagatgaaagcagaaa 862
294  euvalproglyvalleuphellasplivaiahismetleuasplleglu 310
863  ttgttcccggtatattattatcagcgaagtcacatggttgatataaga 912
311  cyspheserlyleuasnrqalaleugluserproleuserproileva 327
913  tgtttttcctttatattatnattggcctttggaaagatgtttccccanct 962
327  lleleualarhnasnarqglyilecysasnvalarglythraspmelt 344
963  catgatgcctacaatagagagatttccaaacacagagacataattaca 1012
344  hrserprohisllyleprovalaspleleuasprtleuvalillele 360
1013  aatpccacatgggttacctgcctgcatttttgatnagctcaattattatt 1062
361  argtnrgluthrtyrlylprothrnglmetilegllleualailear 377
1063  acaactaaaaagttatratgacgacgagattaaacataattttatctatag 1112
377  galaagnvalgluglulleaspmetsplugluserleualaityrleug 394
1113  accacaaagaggaagattgaaactgcattccgatgccttgatcattatga 1162
394  lvgllulegylnglnthrserleuarqhisalaleglnleuilleser 410
1163  ccamaacagstgggaactagtttgcgttgcagacagatatttatctct 1212
411  proalaservalyserythrnsnglyargglulslslecylysai 427
1213  gttgctcacacaattgcattgacagagaaaaaacacactgttgaaatga 1262
427  aaspleuglulvalserclyleuutyrlasपालालyserserlaa 444
1263  agatgfcataaaggcctattatttgcattttttggacagcgcctagatctgta 1312
444  rgleuleuglnglulnglulnargtrylle 454
1313  acgatgttcagagagacgactcacaaatatatt 1344

seq_name: /SIDSL/gcgcdata/geneseq/geneseqn/NA1999.DAT.AAV84507
seq_documentation_block:
ID  AAV84507 standard; DNA; 1167 BP.
XX  AAV84507;
AC
XX
DT  01-MAR-1999 (first entry)

```

DE	Human secreted protein gene 9/ clone n86nzo.	
XX		
KW	Human; secreted protein; fusion protein; gene therapy; protein therapy;	
KW	diagnosis: tissue; cancer; tumour; neurodegenerative disorder; leukaemia;	
KW	developmental abnormality; foetal deficiency; blood; allergy; renal; ds;	
KW	immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;	
KW	inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;	
KW	cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus	
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid digestion;	
KW	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.	
OS	Homo sapiens.	
XX		
PN	W09854963-A2.	
XX		
PD	10-DEC-1998.	
XX		
PF	04-JUN-1998;	98WO-US11422.
XX		
PR	18-DEC-1997;	97US-0070923.
PR	06-JUN-1997;	97US-0048877.
PR	06-JUN-1997;	97US-0048881.
PR	06-JUN-1997;	97US-0048884.
PR	06-JUN-1997;	97US-0048893.
PR	06-JUN-1997;	97US-0048896.
PR	06-JUN-1997;	97US-0048899.
PR	06-JUN-1997;	97US-0048915.
PR	06-JUN-1997;	97US-0048949.
PR	06-JUN-1997;	97US-0048964.
PR	06-JUN-1997;	97US-0049020.
PR	06-JUN-1997;	97US-0049375.
PR	05-SEP-1997;	97US-0057628.
PR	05-SEP-1997;	97US-0057635.
PR	05-SEP-1997;	97US-0057644.
PR	05-SEP-1997;	97US-0057650.
PR	05-SEP-1997;	97US-0057661.
PR	05-SEP-1997;	97US-0057667.
PR	05-SEP-1997;	97US-0057761.
PR	05-SEP-1997;	97US-0057764.
PR	05-SEP-1997;	97US-0057770.
PR	05-SEP-1997;	97US-0057775.
PR	05-SEP-1997;	97US-0057778.
PR	06-JUN-1997;	97US-0048875.
PR	06-JUN-1997;	97US-0048878.
PR	06-JUN-1997;	97US-0048882.
PR	06-JUN-1997;	97US-0048885.
PR	06-JUN-1997;	97US-0048894.
PR	06-JUN-1997;	97US-0048897.
PR	06-JUN-1997;	97US-0048900.
PR	06-JUN-1997;	97US-0048916.
PR	06-JUN-1997;	97US-0048962.
PR	06-JUN-1997;	97US-0048970.
PR	06-JUN-1997;	97US-0048974.
PR	06-JUN-1997;	97US-0049373.
PR	05-SEP-1997;	97US-0057584.
PR	05-SEP-1997;	97US-0057629.
PR	05-SEP-1997;	97US-0057642.
PR	05-SEP-1997;	97US-0057645.
PR	05-SEP-1997;	97US-0057648.
PR	05-SEP-1997;	97US-0057651.
PR	05-SEP-1997;	97US-0057662.
PR	05-SEP-1997;	97US-0057668.
PR	05-SEP-1997;	97US-0057762.
PR	05-SEP-1997;	97US-0057765.
PR	05-SEP-1997;	97US-0057771.
PR	06-JUN-1997;	97US-0048876.
PR	06-JUN-1997;	97US-0048880.
PR	06-JUN-1997;	97US-0048883.
PR	06-JUN-1997;	97US-0048892.
PR	06-JUN-1997;	97US-0048895.

PR 06-JUN-1997; 97US-0048898.  
 PR 06-JUN-1997; 97US-0048901.  
 PR 06-JUN-1997; 97US-0048917.  
 PR 06-JUN-1997; 97US-0048963.  
 PR 06-JUN-1997; 97US-0048971.  
 PR 06-JUN-1997; 97US-0049019.  
 PR 06-JUN-1997; 97US-0049374.  
 PR 05-SEP-1997; 97US-0057627.  
 PR 05-SEP-1997; 97US-0057634.  
 PR 05-SEP-1997; 97US-0057643.  
 PR 05-SEP-1997; 97US-0057646.  
 PR 05-SEP-1997; 97US-0057649.  
 PR 05-SEP-1997; 97US-0057654.  
 PR 05-SEP-1997; 97US-0057666.  
 PR 05-SEP-1997; 97US-0057760.  
 PR 05-SEP-1997; 97US-0057763.  
 PR 05-SEP-1997; 97US-0057769.  
 PR 05-SEP-1997; 97US-0057774.  
 PR 05-SEP-1997; 97US-0057777.

## (HUMA-) HUMAN GENOME SCI INC.

PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,  
 PI Fan P, Feng P, Ferlie AM, Fischer CL, Florence C,  
 PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW,  
 PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Rubin SM,  
 PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;  
 DR WPI; 1999-059865/05.  
 DR P-PSDB; AAM88630.

PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders

Claim 4; Page 358-359; 772pp; English.

XX The invention relates to nucleic acid sequences (AAV84411 to AAV84633)  
 CC encoding human secreted proteins (AAM88534 to AAM88756). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit numbers  
 CC ATCC 97974, 97974, 97975, 97977, 209007, 209008, 209009, 209010,  
 CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host  
 CC cells comprising recombinant vectors containing the nucleic acid  
 CC sequences are used for the recombinant production of the secreted  
 CC proteins. The polynucleotide and amino acid sequences are useful for are  
 CC useful for preventing, treating or ameliorating medical conditions e.g.  
 CC by protein or gene therapy. Pathological conditions can be also  
 CC diagnosed by determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the polynucleotides, based on  
 CC which tissues they are most highly expressed in, and include developing  
 CC products for the diagnosis or treatment of cancer, neurodegenerative  
 CC disorders, developmental abnormalities and foetal deficiencies, blood  
 CC disorders, tumours, leukemias, diseases of the immune system, autoimmune  
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,  
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,  
 CC rstenosis, prostate diseases, obesity, disorders involving osteoclasts  
 CC such as osteoporosis, arthritis or malignancies, diseases of testes/  
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The present sequence represents a gene encoding a human secreted protein  
 CC (see descriptor line for gene number and clone identification).  
 XX

Sequence 1167 BP; 245 A; 332 C; 389 G; 198 T; 3 other;

## alignment\_scores:

Quality: 676.00 Length: 328  
 Ratio: 2.770 Gaps: 5  
 Percent Similarity: 74.380 Percent Identity: 44.207

## alignment\_block:

us-09-589-510-4 x AAV84507

Align seg 1/1 to: AAV84507 from: 1 to: 1167

2 Arg11eGluGluValGlnSerThrSerLysLysGlnArg11eAlaThrH1 18  
 77 AAAGTCCCGAGATCCGATGATTAACAAGATTCAGCAATCGGTGCCA 126  
 18 sThrHisLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleAla 35  
 127 CTCACACATCCGGGACTGGGCTGGACGATGCCCTTGAGACCTGGCAGG 176  
 35 euAlaIaGlyPheValGlyGlnAlaAlaAlaArgGlnAlaGlyLeu 51  
 177 CTCGCAAGGATGCTGGTGCATCGTGGCGACGGCGCGGCTGGCGT 226  
 52 AlaValAspMetIleArgGlnLysLysMetAlaGlyAlaValLeu 68  
 227 GTCTGGAGATGATCCGGGAGGAGATTCGGTGGGCGATCTCTAT 276  
 68 uAlaGlyProProAlaThrGlyLysThrAlaLeuAlaGlyIleAla 85  
 277 TGCTGGCCAGCGCGGACGGGAGAGCGCCATCCCGATGGCATGGCGC 326  
 85 InGluLeuGlySerLysValProPheCysPrometValGlySerGluVal 101  
 327 AGGCCCTGGCCCTGACACGCCATTCACAGCCATCGCGCGCGAGAAATC 376  
 102 TyrSerSerGluValLysLysThrGluValLeuMetGlnAsnPheArg 118  
 377 TTCCTCCGAGATGAGCAACAGCGCGCTGACGAGCGCTCCGCGC 426  
 118 gAlaIleGlyLeuArgIleLysGlnAsnLysGluValTyrGluGlyLuv 135  
 427 GTCCATCGCGCTTCGCATCAAGGAGGAGAGAGATTCAGAGGAGG 476  
 135 aLThrGluLeuSerProGluGluAlaGluSerThrThrGlyTyrAla 151  
 477 TGTGTGAGATCCAGATTGATGACACCAAGGAGGCGCTCCAGGTG 526  
 152 LysSerLieserHisValIleIleSerLeuLysThrValLysGlyThr 168  
 527 GGCAGAA.....CTGACCTTCAAGACCAAGATGAGAC 561  
 168 sGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlyLys 185  
 562 CATCTACGACTGGGACCAAGATGATTGATTCCTGACCAAGACAGG 611  
 185 aAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 201  
 612 TCCAGCGCGGAGCGATCATCGACCAAGGCGGCGGCAAGATCTCC 661  
 202 ArgValAlaGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAla 218  
 662 AAGTGGGCGGCTCTTCACACCGCGCGGAACTGAGCGCTATGGGCT 711  
 218 InGlu.....TyrValProIleProLysGlyLuvAlaHisLysLys 232  
 712 CCAGACCAAGTTCGTCGATGCGCCAGATGGGAGGCTCCAGAAACCAAG 761  
 233 GluIleValGlnAspValThrLeuHisAspLeuAspAlaAlaAlaG1 249  
 762 GAGGTGGTGCACACGTCGTCCTGACGAGATCGACGATCATCACTCG 811  
 249 nProGlnIleGlyLysAspIleLeuSerLeuMetGlyGlnMetMetLysP 266  
 812 CACCCAGGCG.....TTCCTGGCGCTC.....TTCCTG 840  
 266 roArgLysThrGluIleThrGluLysIleuArgGlnGluLysAsnLysVal 282  
 841 GTACACAGGAGGATCAAGTCAAGATTCGTCGAGACATCAATCCCAAG 890  
 283 ValAsnArgTyrTrlLeaspGluGlyIleAlaGluLeuValProGlyValIle 299

```
891 GTGGCTGAGTGGCGGAGGAGGCGAGGAGATCATCCCTGGAGTGT 940
299 uPheIleaspGluValHisMetLeuAspIleGluCysPheSerTyrLeuA 316
|||||
941 GTTCATCGACGAGGTCCACATGCTGGACATCGAGAGGCTTCTCTCTCA 990
316 sNArgAlaLeuGluSerProLeuSerProIle 326
|||||
991 ACCGGGCCCTGGAGAGTGCATGGCGCCTGTC 1022
```

---



[REDACTED]

Dd 361 IKIMLI IPQEMNGI ININX I DO INIOEENENN

heat shock protein

Db 421 GKDSIEKEHVEISELFFYDAKSSAKILADODQKYM 455

RESULT 2

JC5521

TATA-binding protein-interacting protein 49 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Sep-1997 #sequence\_revision 12-Sep-1997 #text\_change 20-Jun-2000

C:Accession: JC5521; PC4474

R:Kanemaki, M.; Makino, Y.; Yoshida, T.; Kishimoto, T.; Koga, A.; Yamamoto, K.; Yamamoto, K.

Biochem. Biophys. Res. Commun. 235, 64-68, 1997

A:Title: Molecular cloning of a rat 49-kDa TBP-interacting protein (TIP49) that is highly

A:Reference number: JC5521; MUID:97339437

A:Accession: JC5521

A:Molecule type: mRNA

A:Residues: 1-456 <KAN1>

A:Cross-references: DDBJ:AB002406; NID:g2225876; PIDN:BA20875.1; PID:g2225877

A:Accession: PC4474

A:Molecule type: Protein

A:Residues: 1-1134-45:129-138;446-452 <KAN2>

C:Comment: This protein plays a role in biological processes such as recombination and

C:Superfamily: conserved hypothetical protein YDR190c

C:Keywords: ATP

F:70-77/Domain: ATP-binding #status predicted <ATP>

F:302-305/Domain: ATP hydrolysis #status predicted <ATP>

Query Match 76.1%; Score 1722.5; DB 2; Length 456;  
Best Local Similarity 72.3%; Pred. No. 1.6e-96;  
Matches 329; Conservative 74; Mismatches 51; Indels 1; Gaps 1;

QY 1 MRIBEVSTSKORATHTHTKIGLDANGALATAAGPVGAAAEAGLAVDMIRQK 60  
Db 1 MKIEVASTTKTQRLASHSHVKGIGLDESGLAQAASGLVGDENAKACGVIVELIKSK 60

QY 61 MAGRAVILAPPGATKATALAGIAQELGSKVPCPMVGESEVSEKKTVEIMENFRAT 120  
Db 61 MAGRAVILAPPGATKATALAGIAQELGSKVPCPMVGESEVSEKKTVEIMENFRAT 120

QY 121 GLRIRKREKVEYEGEVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDAL 180  
Db 121 GLRIRKREKVEYEGEVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDAL 180

QY 121 GLRIRKREKVEYEGEVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDAL 180  
Db 121 GLRIRKREKVEYEGEVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDAL 180

QY 181 IKRVAAGDVITYEANSAGVAVKRGCDSPATEYDLAEAEVPIPKGEVKKKEIVODVTL 240  
Db 181 IKRVAAGDVITYEANSAGVAVKRGCDSPATEYDLAEAEVPIPKGEVKKKEIVODVTL 240

QY 241 HDLDANNAQPOGGODILSLMGOMKPRKTEITKLRQETNKVNNRYIDEGIAELVPGVLF 300  
Db 241 HDLDANNAQPOGGODILSLMGOMKPRKTEITKLRQETNKVNNRYIDEGIAELVPGVLF 300

QY 301 IDEVHMLDIECFSYLNRALESPLSPVILATNRGICNVRG-DMTSPHGIVDLDRLVI 359  
Db 301 IDEVHMLDIECFSYLNRALESPLSPVILATNRGICNVRG-DMTSPHGIVDLDRLVI 359

QY 360 IRTETTYGPEPMIQTALRAQVEIDMEDESLAYIGELGQTSLRHAQLQLSPASVAKTN 419  
Db 360 IRTETTYGPEPMIQTALRAQVEIDMEDESLAYIGELGQTSLRHAQLQLSPASVAKTN 419

QY 420 GREKICKADLEEVSGTLTYDADKSSARLLDQOQERYT 454  
Db 420 GREKICKADLEEVSGTLTYDADKSSARLLDQOQERYT 454

QY 421 GKDSIEKEHVEISELFFYDAKSSAKILADODQKYM 455

RESULT 3

S52698

hypothetical protein YDR190c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YD9346.02c

C:Species: Saccharomyces cerevisiae

C:Date: 19-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 29-Oct-1999

C:Accession: S52698

R:Oliver, K.; Harris, D.

submitted to the EMBL Data Library, March 1995

A:Reference number: S52698

A:Accession: S52698

A:Molecule type: DNA

A:Residues: 1-463 <OLL>

A:Cross-references: EMBL:Z48784; NID:9755782; PIDN:CAA8704.1; PID:g755784; GSPDB:GNO

C:Genetics: A:Gene: MIPS:YDR190c

A:Map position: 4R

C:Superfamily: conserved hypothetical protein YDR190c

Query Match 67.1%; Score 1517.5; DB 2; Length 463;  
Best Local Similarity 65.7%; Pred. No. 3.6e-84;  
Matches 286; Conservative 81; Mismatches 67; Indels 1; Gaps 1;

QY 14 RIATHTHTKIGLDANGALATAAGPVGAAAEAGLAVDMIRQKMGRAVLAPPA 73  
Db 23 RPAATHHTKIGLDGDSGAAKVEGGEVGLDARACGVIVDLIKAKMGRAVLAPPA 82

QY 74 TGTALALGIAQELGSKVPCPMVGESEVSEKKTVEIMENFRATIGLRIRKREYEG 133  
Db 83 TGTALALGIAQELGSKVPCPMVGESEVSEKKTVEIMENFRATIGLRIRKREYEG 142

QY 134 EVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDALIKRVAAGDVITY 193  
Db 143 EVELSPKEASTTGGVAKSISHVITSLKTVGQKLDSSITDALIKRVAAGDVITY 202

QY 194 EANSAGVAVKRGCDSPATEYDLAEAEVPIPKGEVKKKEIVODVTLHDLDANNAQPOGG 253  
Db 203 EANSAGVAVKRGCDSPATEYDLAEAEVPIPKGEVKKKEIVODVTLHDLDANNAQPOGG 262

QY 254 ODILSLMGOMKPRKTEITKLRQETNKVNNRYIDEGIAELVPGVLFIDEVHMLDIECF 313  
Db 263 ODILSLMGOMKPRKTEITKLRQETNKVNNRYIDEGIAELVPGVLFIDEVHMLDIECF 322

QY 314 YLNRALESPLSPVILATNRGICNVRG-DMTSPHGIVDLDRLVIIRTEYGTETIQ 372  
Db 323 YLNRALESPLSPVILATNRGICNVRG-DMTSPHGIVDLDRLVIIRTEYGTETIQ 382

QY 373 ILAIRAQVEIDMEDESLAYIGELGQTSLRHAQLQLSPASVAKTNREKIRKLEEV 432  
Db 383 ILAIRAQVEIDMEDESLAYIGELGQTSLRHAQLQLSPASVAKTNREKIRKLEEV 442

QY 433 SGVLDPADKSSARLLQ 447  
Db 443 KLLFLDARSTKILE 457

RESULT 4

T19534

hypothetical protein C27H6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T19534

R:Gardner, A.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19138

A:Accession: T19534

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-458 <WTL>

A:Cross-references: EMBL:Z81042; PIDN:CAB02793.1; GSPDB:GN00023; CESP:C27H6.2

A:Experimental source: clone C27H6

C:Genetics: A:Gene: CESP:C27H6.2

A:Map position: 5

A:Introns: 22/3; 54/3; 292/1; 413/1; 426/3

C:Superfamily: conserved hypothetical protein YDR190c

Query Match 57.9%; Score 1311; DB 2; Length 458;  
Best Local Similarity 57.8%; Pred. No. 9.7e-72;

[illegible]

OY 419 NREKICKADLEEVSGLYIDAKSSARLLQE 448  
 Db 407 -GGKRVREHVEKAKKEFYADVKRSIAFVEK 435

## RESULT 7

hypothetical protein PH1804 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
 C:Accession: D71191  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: D71191  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-441 <RAW>  
 A:Cross-references: GB:AP000007; MID:93236134; PIDN:BA30923.1; PID:93258240  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH1804  
 C:Superfamily: conserved hypothetical protein YDR190C

## Query Match

Best Local Similarity 43.1%; Score 975; DB 2; Length 441;  
 Matches 207; Conservative 92; Mismatches 131; Indels 18; Gaps 6;

OY 3 IEEVOSTSKORATITHTKIGGLDANGMAIALAAGFVGOAAREAGLAVDMIROKMA 62  
 Db 4 IEEVPAI-KFERGAHSHIRKIGLIDENGKAKFIDGOMGVOKARAGIAVKIKOGSLA 62  
 OY 63 GRAVLLAGPATGKTALALGIAOELGSKVPCPMVGEVSEVSEVKTVMENFRRAIGL 122  
 Db 63 KGIILLVPTGSKTALAMGIAELGSDVPEVOISSEYSEVSEVKTVMENFRRAIGV 122  
 OY 123 RIKENKEVEGEVTELSPEAEESTTGGAKSISHYIISLKTIVKTKOLKDSIYDALI 182  
 Db 123 RISEEKVVEGEVMEKMEIKRTHPNPIEIPESIRITLKKDKKTRIRAGREIAYOLD 182  
 OY 183 EKAVADGVYIYIANGAVKRVGRCSFATVDEAEVVPJPKCEVHKKEIYODVTLAD 242  
 Db 183 MGEVGDVYIQAETGRVSKIG--TTKEEGLEPKKKVVELPTGPVLKIKETVYTLHD 239  
 OY 243 IDAANAPOGGODILSLM--GOMMKPRKTEITEKLRQEVNRYIDEGLAELVPGVLF 300  
 Db 240 LDVNAARAG--IFSLIFGGM-----EINDEIREVNDGTAKOMLEBKRATLVPGVLF 290  
 OY 301 IDEVHMLDIECFSYLNRLSPISPIYIATNKGICINAVRGTDMTSPHGIPVLDLRLVIT 360  
 Db 291 IDECHMLDIEAFSLARMENELAPILILATNRGTRIKRGTIEAPHGIPVMDLRLIIT 350  
 OY 361 RTETVPTMETQIILAIRAOVEIDMEESLAYIGEIQOQSLAHAIOLISPAVSVKTN 420  
 Db 351 NTEPEKKEIETIKIRAKKEVLESEALEYIADLGEKTSRLYAVOLLAPASIIA---G 407  
 OY 421 REKICKADLEEVSGLYIDAKSSARLLQE 448  
 Db 408 GKRVEKEHVEKAREYFADIKRSISFEVK 435

## RESULT 8

T46313  
 hypothetical protein DKFZp434K1011.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: T46313  
 R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Well, B.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000  
 A:Reference number: 223036  
 A:Accession: T46313  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-463 <AAA>  
 A:Cross-references: EMBL:AL136743  
 A:Experimental source: adult testis; clone DKFZp434K1011  
 C:Genetics:  
 A:Note: DKFZp434K1011.1  
 C:Superfamily: conserved hypothetical protein YDR190C

Query Match 43.0%; Score 973; DB 2; Length 463;  
 Best Local Similarity 44.0%; Pred. No. 2,2e-51;  
 Matches 200; Conservative 97; Mismatches 144; Indels 14; Gaps 4;

OY 2 RIEVOSTSKORATITHTKIGGLDANGMAIALAAGFVGOAAREAGLAVDMIROKMA 61  
 Db 9 KVEIRVOTRIERKGAHSHIRKIGLIDALEPPQASOGMGOALARRAAGVLEMIREKTI 68  
 OY 62 GRAVLLAGPATGKTALALGIAOELGSKVPCPMVGEVSEVSEVKTVMENFRRAIG 121  
 Db 69 AGRAVLLAGPGTGKTAIANGMAALGPDPTPTAIGSEIFLEMSKTEALTQAFRRSIG 128  
 OY 122 LRIKENKEVEGEVTELSPEAEESTTGGAKSISHYIISLKTIVKTKOLKDSIYDALI 181  
 Db 129 VRIKEETEIEGEVVEIQRIDPATGTSKVGK----LTKTETMETIYDGTAKMEISLT 183  
 OY 182 KERVAVDGVYIYIANGAVKRVGRCSFATVDEAE--YPIPKGEVHKKEIYODVT 239  
 Db 184 KDVQAGDVITDKATGKISKIGRSTFRADYDAMGSDTKFVQCPDGLQRRKEVHVHS 243  
 OY 240 LHDLDAAAPOGGODILSLMOMMKPRKTEITEKLRQEVNRYIDEGLAELVPGVLF 299  
 Db 244 LHEIDVINSRTG--FLAL----FSGDGEIKSEVRQINAKVAEMREKAKIIPGVLL 296  
 OY 300 FIDEVHMLDIECFSYLNRLSPISPIYIATNKGICINAVRGTDMTSPHGIPVLDLRLV 359  
 Db 297 FIDEVHMLDIECFSYLNRLSPISPIYIATNKGICINAVRGTDMTSPHGIPVLDLRLV 356  
 OY 360 IRTETVPTMETQIILAIRAOVEIDMEESLAYIGEIQOQSLAHAIOLISPAVSVKTN 419  
 Db 357 VSTTPYSEKDTQILIRICEEDVEMSEDATVTLRIGLETSRLYATOLITAAVLGCRKR 416  
 OY 420 GREKICKADLEEVSGLYIDAKSSARLLQEOGERTY 454  
 Db 417 KGEVQVDIKRKYSLFDESRSITQYMKREYODAF 451

## RESULT 9

D69476  
 TBP-interacting protein TTP49 homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
 C:Accession: D69476  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
 G. Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirness, E  
 Nature 390, 364-370, 1997  
 A:Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Arltach, P.; Kaine, B.P.; Sykes,  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: D69476  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-449 <MLE>  
 A:Cross-references: GB:AF000977; GB:AF000782; MID:q2689300; PIDN:AB89434.1; PID:q264  
 C:Superfamily: conserved hypothetical protein YDR190C

## Query Match

41.6%; Score 942.5; DB 2; Length 449;



RESULT 10  
 S61029  
 hypothetical protein YPL235W - Yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: hypothetical protein P1060  
 C:Species: *Saccharomyces cerevisiae*  
 C:date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Jun-2000  
 C:Accession: S61029; S61699; S65260; S65264  
 R:Pohl, T.M.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: S61010  
 A:Accession: S61029  
 A:Molecule type: DNA  
 A:Residues: 1-471 <POH>  
 A:Cross-References: EMBL:Z67751; NID:q1061234; PIDN:CAA91609.1; PID:q1061254  
 R:Uristarazu, L.A.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: S61639  
 A:Accession: S61699  
 A:Molecule type: DNA  
 A:Residues: 1-471 <URR>  
 A:Cross-References: EMBL:X94561; NID:q1181252; PIDN:CAA64252.1; PID:q1181253  
 R:Uristarazu, L.A.; Vissers, S.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S65251  
 A:Accession: S65260  
 A:Molecule type: DNA  
 A:Residues: 1-471 <URF>  
 A:Cross-References: EMBL:Z73591; NID:q1370485; PIDN:CAA97952.1; PID:q1370486; MIPS:YPL2  
 A:Experimental source: strain S288C (AB972)  
 R:Pohl, T.M.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64899  
 A:Accession: S65264  
 A:Molecule type: DNA  
 A:Residues: 1-471 <POW>  
 A:Cross-References: EMBL:Z73591; NID:q1370485; PIDN:CAA97952.1; PID:q1370486; MIPS:YPL2  
 A:Experimental source: strain S288C (AB972)

```

RESULT 11
T46049
RuvB DNA helicase-like protein - Arabidopsis thaliana
N:Alternate names: protein T16K5_180
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence-revision 04-Feb-2000 #text_change 18-Feb-2000
C:Accession: T46049
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemck
Submitted to the Protein Sequence database, January 2000
A:Reference number: 223015
A:Accession: T46049
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <RLE>
A:Cross-references: EMBL:AL132965
A:Experimental source: cultivar Columbia; BAC clone T16K5
C:Genetics:
A:Map position: 3
A:Introns: 251/2
A:Notes: T16K5_180
C:Superfamily: conserved hypothetical protein YDR190c

```

```

QY      1 MREEVOSTSKORIATHTHIKGLDANGMAIALAGFVGQAAREACIAYDMTRQK 60
      2 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4 LRLSEIRDLTRERIGASHIRGLDVSLEPRAVSEGMVGQIKARKAAGVTLELIRDG 63

```

RESULT 12  
E72723  
Probable TATA-binding protein-Interracing protein 49 APE0326 - Aeropyrum pernix (strain C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: E72723  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Kasuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum A:Reference number: A72450; MUID:99310339  
A:Accession: E72723  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1473 <KAM>  
A:Cross-references: DDBJ:AP000059; NID:gs103911; PIDN:BAAT9281.1; PID:gs103965  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0326  
C:Superfamily: conserved hypothetical protein YDR190C

Query Match	38 18;	Score 863;	DB 2;	Length 448;
Best Local Similarity	42.28;	Pred. No. 96-45;		
Matches 191;	Conservative 83;	Mismatches 161;	Indels 18;	Gaps 5
QY	6	VOSTSKORIRATHTHIKGLGDANGMAIALAGFVGQAAREAGLAVDMIROKKMGRA	65	
Db	9	VNDIKVATRVSHSHITGLINLRLEAEVYSGGMQVQAARQAGLLYKMIQDEKINGRA	68	
QY	66	VLLAPRPTGKTALALIAOELGSKYPCFPMYSGEYVSEVAKTIEVLMENFRRAIGLRK	125	
Db	69	LLVTGEPACGTALAIATISKELEDTPFVSIAASEIYNSNEIKTEALTQAFRRMAGIQIK	128	
QY	126	ENKEVEGEVTELSPEBAESTTGGVAKSISHVILSLKTVKSTQOLKDDSYDALIKREKY	185	
Db	129	ETEVELEGEVISTL---EVDRSANGMPKVGK--LIMRTIDMETIYDGLSKMVDACLKREKY	183	
QY	186	AVGVPIYIAENAGVYKVRGCSFAIEYDLEA----EEVYIPRGEVHKKKEIYVDVDTLH	241	
Db	184	MPGVDTQYDAKASGVTRVLGR--SFNSHVDYDMGPKVYLVOCPGGEIQKRRFVHTVCHH	241	
QY	242	DLDAANAPOGODLILSLGMOMKPKRKTETIEKLEPROFINKVVNRYIDEGIAETLVPGVLT	301	
Db	242	DIDVINSRTQYVALES-----GDGTGIEAKAEVADQINNKVLELMRECGAKKVVPPVLT	294	
QY	302	DEVYMLDIEEFSVLNALESPLSPYIATINRGJCNRGDWTMSPHGIPVLDLRLVYIR	361	
Db	295	DEAMMLDIEEFSPLNAIEGELSPLIMATNRLLEKRGTDVDSANGITSDFLDRLLITN	354	
QY	362	TETVGPTEMQILAIRAQVEEIDMOEESLAVLGEIGOOTSRIRAHQILISPAVSSTNKR	421	
Db	355	ALPYTKEDKATILISINCDGEYKIQPTMLDLIVLQATISLRCIFHLIAASEVIRIRSKA	414	
QY	422	EKICKADLEVSGLYIDAKSSARLLIQEOQERYI	454	
Db	415	EIVTVDHISAVRLFEPTKRSEKILTESAGFL	447	

## RESULT 14

T46873

hypothetical protein [imported] - Sulfolobus solfataricus (fragment)

C:Species: Sulfolobus solfataricus

C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 02-Sep-2000

C:Accession: T46873

R:Kulaeva, O.I.; Koonin, E.V.; McDonald, J.P.; Randall, S.K.; Rabinovich, N.; Connaught

Mutat. Res. 357, 245-253, 1996

A:Title: Identification of a DnaB/UmuC homolog in the archaeon Sulfolobus solfataricus.

A:Reference number: 224119; PMID:97030760

A:Accession: T46873

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-112 &lt;KUL&gt;

A:Cross-references: EMBL:052110; PIDN:AAB38088.1

A:Experimental source: ATCC 35091

C:Superfamily: conserved hypothetical protein YDR190C

## Query Match

13.2%; Score 299.5; DB 2; Length 112;

Best Local Similarity 53.2%; Pred. No. 1.2e-11;

Matches 59; Conservative 21; Mismatches 30; Indels 1; Gaps 1;

QY 2 RLEEVQSTKSKORATHTHKGIGDANGMALALAGFVGQAAREAAAGLAVDMIRQKKM 61

Db 3 QIRETKKI-EKEKASINSHITGLDEKGRKAKFIADGLVGVGEAREASGIYVOLIRQKKM 61

QY 62 AGRAVILAGPPATGKTALALGIAOELGSKVPECPMGSEVYSSSEVKKTEYL 112

Db 62 AKGGLFVGPPGPKTALALAIKELGEDPFTTINASEVYSTELKTEYL 112

## RESULT 15

F75154

cell division control protein 48, aaa family (cdc48-1) PAB2086 - Pyrococcus abyssi (stra

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: F75154

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: F75154

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-795 &lt;KAW&gt;

A:Cross-references: GB:AJ248284; GB:AL096836; NID:q5457730; PIDN:CAB49317.1; PID:q545782

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2086

C:Superfamily: transnational endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind

F:530-737/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology &lt;VAT&gt;

## Query Match

8.0%; Score 180.5; DB 2; Length 795;

Best Local Similarity 24.4%; Pred. No. 0.0025;

Matches 91; Conservative 65; Mismatches 132; Indels 85; Gaps 18;

QY 64 RAVLLAGPPATGKTALALGIAOELGSKVPECPMGSEVYSSSEVKKTEV-LMENFRRAIGL 122

Db 217 KGVLLGPPGKTLLAKAVANE--ANAYFIALNGPEIMSKYKYESEERLRLEVFKEA--- 271

QY 123 RIKENKE--VYEGEVELSPEEAESTTGGYAKSISHYIISLKTIVKGTOKLXLDSSIYDAL 180

Db 272 --EENAPAIIFIDIDAIAPKRSE-VTGEVEKRYVAQLLAL--MDGIK----- 314

QY 181 IKEYAVADVYI-----EANSGARVGRGCSFATEYDLAEAEVVP1PKGEVHKKEI 234

Db 315 -----SRGKIVIGATNPPDAIDPALRRPGR-----FDREIEVGVPDRQG---RKEI 358

QY 235 VQDVTIHLDLAANAPOGGODIISLMGOMMKPRK-TEITEKLQELINKVYVNRITIDEGIAE 293

```

Db 359 IQ---IHTRGMPIDPFRKDDVLTKEIDPKRSGKFTKIIDNAIEEVNK-----SKEE 407
QY 294 LVPGLFIDEVHMLDIECFSYLNRALESPL-----SPIVLLATNRGICNR- 339
Db 408 EIQVVL-----KKIDAEIVDEVKTRILDKLDELADVTHGFVGADLALAREAAALALRR 462
QY 340 ----GTDMTSPHGIYPVDLDRVLIITETYPSPTEMIOUAIKRAVEEI-DMDESLAYIG 394
Db 463 LIREGKIDFEAETIIPREVIDELKVTNRDFEALMKVEPSALKREVLIEVNVNMDIGGLE 522
QY 395 EIQGQTSLRHAIO 407
Db 523 DVKOE--LREAVE 533

```

Search completed: November 13, 2001, 06:57:09  
 Job time: 75 sec

Wed Nov 14 08:34:17 2001

us-09-589-510-4.std.rpr

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2001, 06:57:14 ; Search time 11.97 Seconds  
(without alignments)  
1302.109 Million cell updates/second

Title:	US-09-589-510-4
Perfect score:	2263
Sequence:	1 MRIEVQSTSKQRIATH.....YIDAKSSARLIQEQERYIT 455

SCROLLING TABLE: PROPOSED: Gapop 10.0 , Gapext 0.5

```
Searched:      93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	167.5	7.4	903	1	YB56_METUA	Q58556 methanococc
2	154.5	6.8	780	1	SAV_SUUAC	Q07590 synchocyst
3	150.5	6.7	627	1	PTH1_SYNY3	Q55700 synchocyst
4	143	6.3	742	1	CDCH_HAINT	Q9bpf0 halobacteri
5	143	6.3	742	1	CDCH_HAISA	P44664 halobacteri
6	142	6.3	635	1	PTH1_HAINT	P71377 haemophilus
7	140.5	6.2	745	1	VAT_THREAC	Q05209 thermoplas
8	139	6.1	464	1	HSU1_BACHD	Q9ka27 bacillus ha
9	139	6.1	780	1	ARG2_YEAST	P32794 saccharomy
10	135.5	6.0	644	1	PTH1_ODOST	P49825 odontella s
11	135.5	6.0	835	1	CC48_YEAST	P25694 saccharomy
12	134	5.9	628	1	PTH3_PORPU	P51327 pothyria pu
13	133.5	5.9	463	1	HSU1_THEMA	Q9wyw2 thermotoga
14	130	5.7	809	1	TER1_CAEBEL	P54811 caenorhabdi
15	127.5	5.6	760	1	PTH3_CAEBEL	P69492 mycobacteri
16	127	5.6	459	1	HSU1_MYCU	Q9pd94 xyella fas
17	127	5.6	631	1	PTH3_GUITH	Q78516 guillardia
18	127	5.6	1379	1	TBP7_YEAST	P40340 saccharomy
19	126.5	5.6	452	1	P59_MYCU	Q50739 mycobacteri
20	126.5	5.6	837	1	YL34_YEAST	Q07844 saccharomy
21	126	5.6	464	1	HSU1_LACIE	Q48735 lactobacilli
22	124.5	5.5	688	1	DP3X_HAETN	P43746 haemophilus
23	124	5.5	430	1	PRSI_METUA	Q58576 methanococc
24	124	5.5	676	1	YMER_CAEBEL	P34813 caenorhabdi
25	123	5.4	665	1	PTH2_SYNY3	P73179 synchocyst
26	122.5	5.4	398	1	PRSI_ARCFU	Q28030 archaeoglob
27	121.5	5.4	450	1	HSU1_RICPR	Q9zkt8 rickettsia
28	121	5.3	791	1	LON_CAMTE	Q69300 campylobact
29	120.5	5.3	644	1	FTSH_ECOLI	P27891 escherichia
30	120	5.3	747	1	YME1_YEAST	P32799 saccharomy
31	120	5.3	1357	1	RPOB_PSEAE	Q51561 pseudomonas
32	119.5	5.3	444	1	SKDI_HUMAN	Q75351 homo sapien
33	119.5	5.3	695	1	FTSH_LACIA	P46469 lactococcus

34	119	5	3	444	1	SKD1_MOUSE	P46467	mus	musculu
35	118.5	5.2	498	1	FTH3_STREN	O69076	streptococ		
36	118.5	5.2	609	1	FTH3_HERPA	O83746	treponema		
37	118.5	5.2	632	1	FTH3_HELRY	P1408	helicobacte		
38	118	5.2	443	1	HSLU_VITB3	O9kn47	vibrio chol		
39	118	5.2	616	1	FTH4_STEN3	P72991	synchocyst		
40	118	5.2	810	1	TER2_CAEEL	O54813	caenorhabdi		
41	117.5	5.2	410	1	PSR1_METT8	O26824	methanobact		
42	117.5	5.2	440	1	HSLU_PASHA	P32180	pasteurella		
43	117.5	5.2	440	1	PRS4_HUMAN	O35527	homo sapien		
44	117.5	5.2	440	1	PRS4_MOUSE	P49014	mus musculu		
45	117.5	5.2	794	1	SC18_CANAL	P34732	candida alb		

## ALIGNMENTS

```

YB56_METUA 1
ID YB56_METUA STANDARD: PRT; 903 AA.
AC 058556:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CELL DIVISION CYCLE PROTEIN 48 HOMOLOG MJ1156.
GN MJ1156.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96637999; PubMed=86880877;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Adams M.D., Reich C.I.,
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams J.M., Gload A.,
RA Overbeck R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
RA Scott J.L., Geophagen N.S.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Utterback F.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Clifton M.D., Roberts K.M., Smith H.O., Moese C.R., Venter J.C.;
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES. P97/CDC48
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67557; AAB99153.1; -.
CC DR TIGR; MJ1156; -.
CC DR InterPro; IPR001939; -.
CC DR Pfam; PF00004; AAA; 2.
CC DR PROSITE; PS00674; AAA; 2.
CC KW Hypothetical protein; ATP-binding; Repeat; Multigene family.
CC FT NP_BIND 220 227 ATP (POTENTIAL).
CC FT NP_BIND 493 500 ATP (POTENTIAL).
CC SO SEQUENCE 903 AA; 100403 MW; 99F2ABBE2544DC2 CRC64;

```

[illegible]

ID	FTHL_SYNY3	STANDARD;	PRT;	627 AA.
AC	Q55700;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	CELL DIVISION PROTEIN FISH HOMOLOG 1 (EC 3.4.24.-).			
OS	SLR0228.			
OC	Synechocystis sp. (strain PCC 6803).			
OX	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.			
RN	NCBI_Taxid=1148;			
RP	[1]			
RX	SEQUENCE FROM N.A.			
RA	MEDLINE=96127529; PubMed=8590279;			
RA	Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,			
RA	Sugita M., Tabata S.;			
RT	"Sequence analysis of the genome of the unicellular cyanobacterium			
RT	Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb			
RT	region from map positions 64% to 92% of the genome.";			
RL	DNA Res. 2:153-166(1995).			
CC	-1- FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE			
CC	(BY SIMILARITY).			
CC	-1- COFACTOR: BINDS ONE ZINC ION (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41 (ZINC			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D64000; BAI10230.1; -  
DR MEROPS: M41.005; -  
DR InterPro: IPR000130; -  
DR InterPro: IPR000642; -  
DR InterPro: IPR001939; -  
DR Pfam: PF00004; AAA; 1.  
DR Pfam: PF01434; Peptidase\_M41; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE: PS00674; AAA; 1.  
KW Cell division; ATP-binding; Transmembrane; Hydrolyase; Metalloprotease;  
KW Zinc; Multigene family.  
KW DOMAIN 1 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 8 28 POTENTIAL.  
FT DOMAIN 29 117 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 118 138 POTENTIAL.  
FT DOMAIN 139 627 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 212 219 ATP (POTENTIAL).  
FT METAL 433 433 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 434 434 BY SIMILARITY.  
FT METAL 437 437 ZINC (CATALYTIC) (BY SIMILARITY).  
SQ SEQUENCE 627 AA; 68496 MW; 4B2C160213CD0517 CRC64;

Query Match 6.7%; Score 150.5; DB 1; Length 627;  
Best Local Similarity 21.0%; Pred. No. 0.051;  
Matches 105; Conservative 75; Mismatches 159; Indels 161; Gaps 26;

QY 25 GLDNGMAIALAAGFVGQAARPAAGLAADMIRQKMAAGRAVILAPPTGKTALALGTA 84  
DB 176 GIDAKKELEOVYFELKOPERTAVGAKIP- - - - - KGVLLVPPCTGKTLAKAIA 226  
QY 85 QELSGKYPCPMVSGE- - - - - VSEYEVKTEVLEMFRAIGLRIKENKE- - - - - VYEGEV 135  
DB 227 GEAG- - - - - VPFSISGESEVEFVGASRVR- - - - - DLFKKA- - - - - KENAPCLIFIDBI 273  
QY 136 TELSPREASTTGG- - - - - YAKSISHVILSLKTVKQOLKIDSSITDALIKEKVAQDVLY 192  
DB 274 DAGVQRGAGIGGNDREDEQTLNOLTEMDFEGN- - - - - TGIIT 314  
QY 193 IEAN- - - - - SGAVKRVGRCSFATEYDLAEVVPPIPKG- - - - - EVH- - - - - KKEIVQDVTLH 241  
DB 315 AATNPDPVLDALMRPGR- - - - - FDRQVWADAPDY- - - - - SGKKEILEVYARKKTLAPVSI- 366  
QY 242 DLDAANAPOGODILSLMGQ- - - - - MMKPKTEITELKROELINKVNRITD- - - - - EG- 290  
DB 367 DSIAKRTGFGAGDLANLNEALITLARRKSAIT- - - - - LEIIDAVIDRVVAGMEGPPLVD 423  
QY 291 - - - - - IAEVPGVLFIDEVHMLDIECFSYLNRALESPLSVIATNRGICNVR 339  
DB 424 SKSKRLIAYHEVGHAIIVGTLLKD- - - - - HDVQKTLIP- - - - - RG- - - - - QAQ 462  
QY 340 GTDMSPHGIPVDLDRIVITRTETVPTMIQILAI- - - - - RAQVEIDMDEE- - - - - 388  
DB 463 GLTWTFPN- - - - - EROGLITKAOLMARITAGAMGRAAEVEFVGDDVTTG 507  
QY 389 - - - - - SLAVIGETGOQTSLRHAIDILSPASVSKTN- - - - - GRKIKKADLEEVSGIYLDA- 439  
DB 508 AGCDLQGVTEMAROWVTRGMSNIGPISLESSGGEVFLGGILMNSSEEVATRIDAOV 567  
QY 440 - - - - - KSSARLLOEOE 451  
DB 568 RQLAEQGHOMAKIYDEORE 587

RESULT 4  
CDCH\_HALNI STANDARD; PRT; 742 AA.  
AC Q9HBF0;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE CDCH PROTEIN  
GN CDCH OR CDC48C OR VNG1667G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium.  
OX NCBI\_TaxID=64091;

RA [1]  
RP MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahias G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Shroga J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Jung K.-H.,  
RA Isenbarger T.A., Peck R.F., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Alam M., Freltas T., Hou S., Daniels C.J., Hood L., Dassarma S.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Wood L.,  
RT "Genome sequence of Halobacterium species NRC-1."  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
CC -1- FUNCTION: MAY BE PART OF A TRANSDUCTION PATHWAY CONNECTING LIGHT  
CC -1- TO CELL DIVISION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES. P97/CDC48  
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE005075; AAG19919.1; -  
DR PROSITE: PS00674; AAA; 2.  
KW ATP-binding; Repeat.  
FT NP\_BIND 230 237 ATP (POTENTIAL).  
FT NP\_BIND 503 510 ATP (POTENTIAL).  
SQ SEQUENCE 742 AA; 81657 MW; 4E248D2P87ABE126 CRC64;

Query Match 6.3%; Score 143; DB 1; Length 742;  
Best Local Similarity 19.6%; Pred. No. 0.18;  
Matches 90; Conservative 74; Mismatches 151; Indels 144; Gaps 18;

QY 15 IATHTRKIGIGLDANGMAIAL- - - - - AAGFV- - - - - GQAARPAAGLAADMIRQKMA- - - - - 62  
DB 151 IAVETPEGVCLVIEDVDVLEPREPISGFERTGGITIEDIGLENIQVRKVELPK 210  
QY 63 - - - - - GRAVILAPPAATGKTALALGIAOELGSKVPCPMVSGSEVSEVKKTE 110  
DB 211 HPQIFQKIGLEPPQSVLLHPPGTGKTLAKAVANE- - - - - TSASFPSIAGPILISKYGESE 268  
QY 111 - - - - - VLMEFRAIGLRIKENKE- - - - - VYEGEVTLSPEASTTGGYAKSISHVILSLKTVKKT 167  
DB 269 QQLRELFEDA- - - - - KDSPSILFIDELDSIAPRREVDGTEVERRVVQAQLTMMQDIEGR 323  
QY 168 KQIKLDSITDILKEKAVAGDVITYEANSAGVAVRGCSFATEYDLAEVVPPIPKG- 226  
DB 324 GOV- - - - - VTAATNRVDAVDALRRPGR- - - - - FPRELEIGVPRDITGR 362  
QY 227 - - - - - EVNKK- - - - - KEIVQDVTLHDIDAANAPOGODILSLMGOMMKPKTEITELKROELINK 281  
DB 363 EELIKHTHGMPLSDDVNLSTL- - - - - ADDTGFGVADIESLS- - - - - KEAAMR 406  
QY 282 VNRVYIDEGIAELVPGVLFIDEVHMLDIECFSYLNRALESPLSVIATNRGICNVRGT 341

```

Query Match      6.38; Score 143; DB 1; Length 742;
Best Local Similarity 19.68; Pred. No. 0.18;
Matches 90; Conservative 74; Mismatches 11; Indels 144; Gaps 18;

QY 15 IATHTHKGGLDANGMALN-----AAGFV-----GQAARAAGLAVDMIRKMA----- 62
Db 131 IAVTEPEGCVLCTEDTDVLEEREPTSGFERTGGITVEDIGGLENEIGQVREAVELPMK 210
QY 63 -----GRAVLGAPPATGTALAIAGIDGSKVPCFPMNGSVYSSEVKTE 110
Db 211 HPOIFOKLTGIEPQGVLLGHGPGCTKTLAKRVANE--TSSFSIAGPELISKYGESE 268
QY 111 -VLMEENRRRAIGLRKENKE--VYGEVYELSPBEASTGTGCAKSIASHVYISLKTAKCT 167
Db 269 QQLREIFEDA-----KDDSPSITLFIDELDDSTAPKREYVTEVEVERRYVAQLLTIMMDSLEGR 323
QY 168 KQLKLDSSIYDALIKEKAVAGDVIIYEAANSQAVRVGRCSFATEYDLAEAEVPIPKS- 226
Db 324 GQV-----IYIATNREVDVADVDALRRPR-----PDRELEIGVDPDEIGR 362
QY 227 ----EYHKK-KEIVQDVTLLHDDAANQPOGGQDIISLGMQMKPKPKTETTEKLROEINK 281

```

```

RESULT 6
FTHL_HAELN ID FTHL_HAELN STANDARD; PRT; 635 AA.
AC P71377;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CELL DIVISION PROTEIN FTHS HOMOLOG 1 (EC 3.4.24.-).
GN FTHS-A OR HT1335.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
ON NCBI_TaxId=727;
RX [1]
SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shilly R., Liu L.-I., Glode A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ullrich L.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
RN [2]
RP REVISIONS.
RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.;
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE
CC (BY SIMILARITY).
CC -! COFACTOR: BINDS ONE ZINC ION (POTENTIAL).
CC -! SUBCELLULAR LOCATION: INTERAL MEMBRANE PROTEIN.
CC -! SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41 (ZINC
CC METALLOPROTEASE).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC CC or send an email to license@isb-slb.ch).
CC CC -----
DR EMBL: U32812; AAC22979.1; -.
DR TIGR: H11335; -.
DR InterPro: IPR000642; -.
DR InterPro: IPR001939; -.
DR Pfam: PF00004; AA; 1.
DR Pfam: PF01434; Peptidase_M41; 1.
DR PROSITE: PS00674; AA; 1.
KW Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;

```



KM Zinc. 1 4 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1 24 POTENTIAL.  
 FT TRANSMEM 25 95 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 25 120 POTENTIAL.  
 FT DOMAIN 121 635 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 191 198 ATP (POTENTIAL).  
 FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 414 414 BY SIMILARITY.  
 FT METAL 417 417 ZINC (CATALYTIC) (BY SIMILARITY).  
 SO SEQUENCE 635 AA; 70030 MW; 70E6DD1E8293839D CRC64;

Query Match Best Local Similarity 23.9%; Score 142; DB 1; Length 635;  
 Matches 85; Conservative 51; Mismatches 116; Indels 104; Gaps 21;

OY 37 AGFVQAARERAGLAVDMIRQ---KKMG---RAVLGAPPATGKTLALGIAOELGS 89  
 DB 151 ADVAGCEAKKEVEIVDFLDPKFNOLGKIPKGLMVGPGTGKTLARAIAGE--A 208  
 OY 90 KVPCCPMVGSB-----VYSEVKTEVLMENFRRAIGRIKENKE--VYEEVTELS 140  
 DB 209 KVPFTTISGDFVEMFVGASRYR-----DMEEQA---KKNAPCLIFIDEIDAVGR 257  
 OY 141 BEAESTTGCT---AKSISHYISLKYKTKKOLKLDSSYDALIKEKVAAGVYIEANS 197  
 DB 258 ORGAGLGGHDEROTLNMIVEMDGFSGN---DSVIYIAATNRP---DVL-----D 303  
 OY 198 GAVKRVGCDSPATEYDELEAEVYIPRG---EVHKRR-ETVODYTLHDLDAANAPO 251  
 DB 304 PALTRPGR-----FDROVVGGLPDYKRGQILKVMHRKVSADVDAMTL-ARGTPG 355  
 OY 252 GGDILSLMGQ---MMKPKTEIT---EKLROEIN-----KVNRYIDE 289  
 DB 356 SGADILNVEALPAARANKRTVMTLEFEKAKDKINKGPRKRMIMTDKQESTATHEA 415  
 OY 290 G---TAEIAPGVLFIDEVHMDI-----ECFSYLNRALSPLSPT 326  
 DB 416 GRAIVGYLVE---HDPVHKVITIIIPRGALGVTFPLPEGDISISOKLESKSTL 468  
 RESULT 7  
 VAT\_THEAC STANDARD; PRT; 745 AA.  
 -AC 005209;  
 DT 01-OCT-2000 (rel. 40, Created)  
 DT 01-OCT-2000 (rel. 40, Last sequence update)  
 DT 01-OCT-2000 (rel. 40, Last annotation update)  
 DE VCP-LIKE ATPASE.  
 GN VAT OR TA0840.  
 OS Thermoplasma acidophilum.  
 OC Archaea: Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;  
 OC Thermoplasma  
 OX NCBI\_TaxID=2303;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RC MEDLINE=97227958; PubMed=9119075;  
 RA Baumeister W., Tamura T., Lupas A.N., Peters J., Cejka Z., Ashraf W.,  
 "Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase  
 homologue from the archaeon Thermoplasma acidophilum".  
 RL FEBS Lett. 404:263-268(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RC MEDLINE=20479972; PubMed=11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Korte K.K., Volker C.,  
 Mewes H.-W., Fritsman D., Stocker S., Lupas A.N., Baumeister W.,  
 "The genome sequence of the thermophilic scavenger Thermoplasma  
 acidophilum".  
 RL Nature 407:508-513(2000).

RN [3]  
 RP 3D-STRUCTURE BY ELECTRON TOMOGRAPHY.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE=99283870; PubMed=10356978;  
 RA Kockel B., Walz J., Hegert R., Peters J., Typke D., Baumeister W.,  
 "Structure of VAT, a CDC48/p97 ATPase homologue from the archaeon  
 Thermoplasma acidophilum as studied by electron tomography".  
 RL FEBS Lett. 451:27-32(1999).  
 CC -1- FUNCTION: HAS AN OPTIMAL MG2(+)-ATPASE ACTIVITY AT 70 DEGREES  
 CC CELSIUS.  
 CC -1- SUBUNIT: HOMODIMER. FORMS A RING-SHAPED PARTICLE.  
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES. P97/CDC48  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch))  
 CC  
 CC EMBL: U78072; AAC45089.1;  
 CC DR EMBL: AL445065; CAC11969.1;  
 CC DR InterPro: IPR001939;  
 CC DR Pfam: PF00004; AAA; 2.  
 CC DR PROSITE: PS00674; AAA; 2.  
 CC NP\_BIND: 231  
 CC ATP-binding: Repeat. 238  
 CC NP\_BIND: 508 515 ATP (POTENTIAL).  
 FT SEQUENCE 745 AA; 83138 MW; B195EB2044FCF1F2 CRC64;

Query Match Best Local Similarity 21.3%; Score 140.5; DB 1; Length 745;  
 Matches 93; Conservative 70; Mismatches 148; Indels 125; Gaps 21;

OY 64 RAVLLGAPPATGKTLALGIAOELGSVYPCPMVGSSEYV-----SEVKTEVLMENFR 118  
 DB 225 KGVILXPGPGTKTLARAVANESGAN--FLSINGPELMSKYGSSOKLRELIPSKAEFT 282  
 OY 119 AIGLRKENKEVEVEVTELSPEAEETGGYAKSISHVILSTVYKTKOLKLDSSYD 178  
 DB 283 APSI-----FIDEIDSIAPKREE-----VOGEVRRVAQLLT 316  
 OY 179 AL--IKER---VAVGDVYIEANSQAVKRGCDSPATEYDELEAEVYIPKGEVHKKE 233  
 DB 317 LMGKMERGHVIVIGATNRIIDAIDPALRPRG-----FDREIEIGVDRNG---RKE 365  
 OY 234 IVODYTLHDLDAANAPOG-----GODILSLMGQ----- 262  
 DB 366 ILM---IH---TRNPLDMSSEKKELEEMADYTFGVGADLALVRESAMNLRRL 418  
 OY 263 ---MMKPKTEITEKL-----ROEINKVYNYIDEGIAELVPGVLFIDEVHMDIEC 311  
 DB 419 PEIDLDKPIPTELLEKMVYTEDDFKNAKLSIEPSSIREYVWE--VNVYHMDIDGLEDK- 476  
 OY 312 FSYLNRALESP-LSPIYIATNNGICNVGDTMSPHGIPVDLRLVIRTEY-----Y 365  
 DB 477 -REIKETVELPLKPDVE--KRUGIRSPKGFELYGPPGKTKLLKAVATESNANPISIK 533  
 OY 366 GPFEMIGIL-----AIRAQVEEIDMEESLAYGEIGOOTSLRHAQIOLISPAVSYSKING 420  
 DB 534 GPEVLSKMGVSEKAHEIRKAKOVAPALVFLDEIDS-----IAPRRGTTSDSG 583  
 OY 421 -REKICRADLEEVSG 435  
 DB 584 VTERIVQLLTSLDGI 599  
 RESULT 8  
 HSDU\_BACHD STANDARD; PRT; 464 AA.  
 ID HSDU\_BACHD

AC G9KA27;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ATP-DEPENDENT HSL PROTEINASE ATP-BINDING SUBUNIT HSLU  
 GN HSLU OR CLP1 OR BH2463.  
 OS Bacillus halodurans.  
 OC Bacillus/Staphylococcus group; Bacillus.  
 NCBI\_TaxID=86665;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- FUNCTION: CHAPERONE SUBUNIT OF A PROTEASOME-LIKE DEGRADATION  
 CC COMPLEX (BY SIMILARITY).  
 CC -1- SUBUNIT: INTERACTS WITH HSLV (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOSOL (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY. HSLU SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: AB001515; BAB06182.1;  
 KW Chaperone; ATP-binding.  
 FT NP\_BIND 58 65  
 SQ SEQUENCE 464 AA; 52634 MW; 6575FED05823181 CRC64;

Query Match Best Local Similarity 6.1%; Score 139; DB 1; Length 464;  
 Matches 84; Conservative 76; Mismatches 122; Indels 126; Gaps 23;

QY 40 VGOAAREAGLAV-----DMROKMGAGRAVLAPATGKTALALGIAOELGS 89  
 DB 19 VGOEAGKRSVAIALNNRRRPOLEKLD-EITPKNIMIGTGVKTEIARLAKLVGA 77  
 QY 90 KYPECPM-----VGSEVYS--SEVKTEVLMENFRAIGLRKKEKVEYEGEYVE 137  
 DB 78 --PFWKVEATKFTVEGYGRVDSMINDIVETSVRLVKEEKMGVKKOAEQANORIVEL 135  
 QY 138 LSPFEASTGGYAKSISHVIISLTKYKTKOLKLD-----SJIYDAL-----IKKRVAV 187  
 DB 136 LVP--AKKKTAYKN-----PLEMLFGOOOEWTDHQETSAPQDPLTERRRKMSQDLAL 186  
 QY 188 GDVIYIANSAGVAVRGDSFATEYDLEAEVYPIPKGVHKKELVODVTLHIDDAAN 247  
 DB 187 GELE-----DHYVT--VEVEEOTP-----OFFDMLQGS 212  
 QY 248 AOPGODILSLMGOM--KPKKTEITTEKL--OEINKVNVRYIDEIGIAELVPP-- 296  
 DB 213 GMEOMGMNMGEMGNMMPKRRKKRKLTVREARVLTVEEEOKLD--MEVYOEVSAAE 270  
 QY 297 --GVLTIDEVHMD-----DIECSYLNRALESPLSIYVILATNRGICNVRGDTMSP 346  
 DB 271 OLGIIVFIDEIDKTAGGOOSADVS-----REGVORDIPI-----VEGSTVYTK 314  
 QY 347 HGIVPDLRLVITRTETV--GPTEMIOIL-----AIRAOVEIDMDE 387  
 DB 315 YG-PVS-TDHMLFAGAHVAKPSPDLPELGRRPIVELSGLNLTVD 360

RESULT 9  
 ID AFG2\_YEAST STANDARD; PRT; 780 AA.  
 AC P32794;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE AFG2 PROTEIN.  
 GN AFG2 OR YLR397C OR L8084.16.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94152174; PubMed=8109176;  
 RA Thorsness P.E., White K.H., Ong W.-C.;  
 RT "AFG2, an essential gene in yeast, encodes a new member of the  
 RT Sec18p, Pas1p, Cdc48p, Tbp-1 family of putative Atfases."  
 RL Yeast 9:1267-1271(1993).  
 CC [2]  
 CC SQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
 RA Favell A., Fulton L., Gallung S., Greco T., Kirsten J.,  
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,  
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,  
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,  
 RA Riffen L., Riles L., Tatch A., Trevisan E., Vignati D.,  
 RA Wilcox L., Wolfdman P., Vaudin M., Wilson R., Waterston R.;  
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: NOT YET KNOWN. ESSENTIAL FOR VIABILITY.  
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: L14615; AAC37367.1;  
 DR EMBL: U19729; AAB82355.1;  
 DR PIR: S39110; S39110.  
 DR SGD: S0004389; AFG2.  
 DR InterPro: IPR001939;  
 DR Pfam: PF00004; AAA; 2.  
 DR PROSITE: PS00674; AAA; 2.  
 DR NP\_BIND 286 293  
 FT NP\_BIND 557 564  
 SQ SEQUENCE 780 AA; 84747 MW; 75094DFA30401D4E CRC64;

Query Match Best Local Similarity 6.1%; Score 139; DB 1; Length 780;  
 Matches 62; Conservative 44; Mismatches 89; Indels 60; Gaps 12;

QY 62 AGRAVLLAGPPATGATALALGIAOELGSVPPGPMGVSYSSEVKTE-VLMENFERRAI 120  
 DB 549 APKGVLLYGPGRGSKTLTKALATSG--INFLAVKGPETFKYGESERAIREIFKRA- 605  
 QY 121 GLAIKKEKVEYEGEVELSPEEAESTTGGYAKSISHVIIS-LKTVGKTKOLKDSIYDA 179  
 DB 606 --RSAAPSIIFFEDIDALSPDRDGSST--SAANHVLISLNLNEDGYEELK----- 651  
 QY 180 LIKEKVAVDVYIERN-----SGAVKVRGDSFATYDLEAEVYPIPKGVHKKKEI 234  
 DB 652 -----GVTVAAATNPREDIDALLRPRGLD-----RHIVGPPVPPNARLEI 692  
 QY 235 VDDVT-----LHDLDAANAOPGODILSLMGOMKPKKTEITTEKLNOEINKVY 283

117 RRAIGLRIKENE - VÆGEVELSPLEASTING - IANSISHTILSKENTVANGEN 172

DR PFam:PF00004; AAA; 2.  
DR PROSITE: PS00674; AAA; 2.  
KW Cell cycle; ATP-binding; Repeat.  
FT NP-BIND 255 362 ATP (POTENTIAL).  
FT NP-BIND 526 533 ATP (POTENTIAL).  
SQ SEQUENCE 835 AA; 91995 MW; 02ADB8A9A227614D8 CRC64;

RESULT	ID	FTSH_PORPU	STANDARD:	PRT:	628 AA.
12	FTSH_PORPU				
AC	P51327;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	CELL DIVISION PROTEIN FTSH HOMOLOG (EC 3.4.24.-).				
GN	FTSH OR YCP25.				
OS	Porphyrin purpurea.				
OG	Chloroplast.				
OC	Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyrin.				
ON	NCBI-TaxId=2787;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RT	STRAIN=AVONPORT.				
RA	Reich M.E., Munholland J.:				
RT	"Complete nucleotide sequence of the Porphyrin purpurea chloroplast				
RT	genome."				
RL	Plant Mol. Biol. Rep. 13:333-335(1995).				
CC	-I- FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE				
CC	(BY SIMILARITY).				
CC	-I- COFACTOR: BINDS ONE ZINC ION (POTENTIAL).				
CC	-I- SUBCELLULAR LOCATION: CHLOROPLAST; INTEGRAL MEMBRANE PROTEIN				
CC	(POTENTIAL).				
CC	-I- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.				
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.1 (ZINC				
CC	METALLOPROTEASE).				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).				
CC					
DR	EMBL: U38804; AAC08213.1; -.				
DR	MEROPS: M41.005; -.				
DR	InterPro: IPR000130; -.				
DR	InterPro: IPR000642; -.				
DR	InterPro: IPR01939; -.				
DR	Pfam: PF00004; AAA; 1.				
DR	Pfam: PF01434; Peptidase_M41; 1.				
DR	PROSITE: PS00674; AAA; 1.				
DR	PROSITE: PS00142; ZINC_PROTEASE; 1.				
KW	Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;				
KW	Zinc; Chloroplast.				
FT	TRANSMEM	8	28	POTENTIAL.	
FT	TRANSMEM	119	139	POTENTIAL.	
FT	NP_BIND	213	220	ATP (POTENTIAL).	
FT	METAL	434	434	ZINC (CATALYTIC) (BY SIMILARITY).	

QY	39	FVGQAAAEAGLAWDMIRKKMAGRVVLVAGPAPCTKTLALGIAQELGSKVPCPCPMWG	98
Db	191	FLKQPESTVAAGKIP-----KGVLLGPGPGTIGTLAALIGEAG--VPFISIG	239
QY	99	SE-----VYSSEKTEVLMENFRRAIGLRIKENK--VYEGEVELSPPEASTGG	149
Db	240	SEFVEVMEVGVASHVR-----DLFKKA-----KDNAPCIVFIDEIDVNGRQGTGVGG	288
QY	150	---YAKSISSHIIISLKVKGKTKQLKLDSSITYDALIKEVAVGVITYEAN----SGAVK	201
Db	269	NDEBQDTLNLTLMDFEGN-----TGVIYAATNRADILDSALL	329
QY	202	RVGRCDSFATEYDLEAEVYPIPKG-----EVH-KKKEIVQDVTLDLDAANAQOQGGD	255
Db	330	RPRG-----FDRQVSDVPDPRGRALILEVAAKKKMKESKVSLETT-ARRTPGFSGAD	381
QY	256	ILSLMGQ---MMKPKTEITEKLRQEKINKVNRVYIDEGIAELVGVLFIDEVHMLD---	308
Db	382	LANLINALTAALITARRRKSAMTWS-----EIDTSIDRVAVG---LEGTPLIDSKS	427
QY	309	--IECFEYLRAI-----ESPLSPIVLATNRGICANRGTDMTSPGIPVDLLRLVT	359
Db	428	KRLIAIVEVGHATIGSLLEHHDVPQVTLIP--KG--QARGLTWTPSPD-----DQSLI	477
QY	360	ITETVGPTEMIQIILIRAQVEIIDDE-----SLAVLEIGQDTSLRRAIQLISPA	412
Db	478	SRSQIL--ARIVGALGSRAAELIEDDAVTTGASNDLQOYTSMAKQVYTRFGMSKIGPL	535
QY	413	SVYSKIN---GR-----EKICKADLEVSGLYLDAK	440
Db	536	SLESQGSDFPLGRMGSGSEYSDDEVATNIDKQREIVSECYKEAK	580
RESULT	13		
HSLU	THEMA		
ID	HSLU	THEMA	STANDARD: PRT; 463 AA.
AC	Q9WY22;		
DT	01-OCT-2000	(Rel. 40, Created)	
DT	01-OCT-2000	(Rel. 40, Last sequence update)	
DT	01-OCT-2000	(Rel. 40, Last annotation update)	
DE	ATP-DEPENDENT HSL PROTEASE	ATP-BINDING SUBUNIT	HSLU.
GN	HSLU OR TM0522.		
OS	Thermotoga maritima.		
OX	Bacteria; Thermotogales; Thermotoga.		
NCBI_TaxID	2336;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MSB8 / DSM 3109.		
RX	MEDLINE=99287316; PubMed=10360571;		
RA	Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,		
RA	Halt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,		
RA	McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,		
RA	Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,		
RA	Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,		
RA	Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;		
RT	"Evidence for lateral gene transfer between Archaea and Bacteria from		
RL	genome sequence of Thermotoga maritima."		
RL	Nature 399:323-329(1999).		
CC	-1- FUNCTION: CHAPERONE SUBUNIT OF A PROTEASOME-LIKE DEGRADATION		
CC	COMPLEX (BY SIMILARITY).		
CC	-1- SUBUNIT: INTERACTS WITH HSLU (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE CLP CHAPERONE FAMILY. HSLU SUBFAMILY.		

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed, usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC			
DR	EMBL: AE001728; AAD55607.1; -.		
DR	TIGR: TM0522; -.		
KW	Chaperone; ATP-binding.		
FT	NP_BIND 60 67	ATP (POTENTIAL).	
SQ	SEQUENCE 463 AA; 53052 MW; F871CD909CFE5A5CA CRC64;		

Query Match 5.9%; Score 133.5; DB 1; Length 463;  
 Best Local Similarity 20.4%; Pred. No. 0.36;  
 Matches 99; Conservative 66; Mismatches 117; Indels 203; Gaps 24

[illegible]

```

Qy 151 -AKSISHVI-----ISLTKVAGTK-----QKLDSSLYDALIRKAVAGDT--YIE 194
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 95 VGNKDMSRI RDLVEISVNNVQKQKEVERQAEELVEERILDLAPESKAMPVYTNFIN 154
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 195 ANSGAVKRVGRCDSPATEYDLAEELVYP-----IPGGEVHKKK--EIV 235
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
      155 195C-----00000YTPEDRRRPRAKREERKIRKGLGEDELEDELE 198

```

```

06      133  L1186      *-----*
QY      236  QDVT-----LHDLDRANKPOGQDILSLKQMM-----KPKTEITEK-----PROE 278
      : : : : :
199  ETVSPFMGJFGQMEDL-----GLETNMGSLPKRRKKRKKMSARAVLLPLE 249
Db
QY      279  INVVVRITDEGIAELVP-----GVLEIDVEMLDIECFSYLNRALESPLIVLATNR 333
      : : : : :

```

```

Db 250 AELIID-MDKVVOELMDRAQNRGIIIFIDEIDKI-----AGKESAVPDV-----29
QY 334 GICNVKGTDMTSPHGIPVDDLRLVIRRET-----YEPPEMQLIAIAQVEIIMDEES 38
293 -----SROGOIORDILP-----IVEGTTIMTKGVPRTIDFIIA-----32
Db 390 IAYVIGIGOOTSLRHAIOLISPAVSVYKNGREKICKADLEVSGLYLDANSSARLLIQEO 44

```

```

QY 328 -----GAFHVSRPDDLPELQGRPI RVELSP-----LTEEDFVRIKEP 367
QY 450 QERYI 454
QY : |
Db 368 ENAII 372

```

	RESULT	14
TER1_CAEEL		
ID	TER1_CAEEL	STANDARD: PRT; 809 AA.
AC	P54811;	
DT	01-OCT-1996 (rel. 34, Created)	
DT	01-OCT-1996 (rel. 34, Last sequence update)	
DT	01-OCT-1996 (rel. 34, Last annotation update)	
DE	TRANSITIONAL ENDOPLASMIC RETICULUM ATPASE HOMOLOG 1 (P97/CDC48 HOMOLOG 1)	
DE	1)	
GN	C06A1.1.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	

RA McMurray A.: Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RL -I- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES. P97/CD448  
 CC SUBFAMILY.  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its use  
CC by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL: Z49886; CAA90050.1; -
DR	Wormped: C06AL.1; CE02114.
DR	InterPro: IPR01939; -
DR	Pfam: PF00004; AAA, 2.
DR	PROSITE: PS00674; AAA, 1.
KW	Hypothetical protein; ATP-binding; Repeat; Multigene family.
FT	NP_BIND 251 258 ATP (POTENTIAL).
FT	NP_BIND 524 531 ATP (POTENTIAL).
SQ	SEQUENCE 809 AA; 89785 MW; ECF02EFB939FA777 CRC64;

Query Match	5.7%;	Score 130;	DB 1;	Length 809;
Best Local Similarity	25.9%;	Pred. No. 1,2;		
Matches	59;	Conservative	38;	Mismatches 93;
				Indels 38;
				Gaps 11

```

QY      64  RAVLLAGPATKITALAIGIAOELCKAPPCPMNGSEVYSSEVAKTEVLEMNFRATIGR 123
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      245  RGLLEPGPGCTKTLIARAVANETSF--FLLINPEVMSKSGESE---SNLRKPE-E 299
QY      124  IKENKE--AYGEVVELSEPEASTGTGYAKSI-SHVILISLTKVGTOKLUDSIYDAL 186
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      300  ANNOGALITREFFITALAKR--FKINGVEFRIYVOLLTLMDGVGSRNL----- 347

```

[illegible]

```

Db      397  ADDVDLEQI ANECHGVGADLASLQSEALQIRKRMELIDLEDQI 444
RESULT  15
FTSH_MYCTU
ID      FTSH_MYCTU
AC      R06947.
NC

```

AC 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 36, Last annotation update)  
 DE CELL GROWTH PROTEIN FTSH HOMOLOG (EC 3.4.24.-).  
 GN FFSH OR RV3610C OR MTC07H7B.12.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 Actinomycetia; Actinomycetiales; Actinomycetaceae; Actinomy-

OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=773;  
PN []  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MERLIN=88295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Church C., Harris D., Gordon S.V., Eigmler K., Gass S., Barry C.E. III, Tekle A.F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor K.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jacobs K., Krog A., Mclean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,  
RA Taylor K., Whitehead K., Barrett B.G.;  
RT "Deciphering the biology of *Mycobacterium tuberculosis* from the  
RT complete genome sequence." ;  
RL Nature 393:537-544(1998).

```

RN [2]
RP SEQUENCE OF 203-323 FROM N.A.
RC STRAIN-H37RA;
RX MEDLINE=98391006; PubMed=9729123;
RA Anilkumar G., Chauhan M.M., Ajitkumar P.;
RT *Cloning and expression of the gene coding for FtsH protease from
RT Mycobacterium tuberculosis H37Rv.";
RL Gene 214:7-11(1998).
CC -!- FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE
CC (BY SIMILARITY).
CC -!- COFACTOR: BINDS ONE ZINC ION (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41 (ZINC
CC METALLOPROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z95557; CAB08956.1; -
DR EMBL; Z83338; CAB05953.1; -
DR MEROPS; M41.001; -
DR TubercuList; RV3610C; -
DR InterPro; IPR000642; -
DR InterPro; IPR001939; -
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF01434; Peptidase_M41; 1.
DR PROSITE; PS00674; AAA; 1.
KW Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;
KW Zinc
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 POTENTIAL.
FT DOMAIN 27 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 131 POTENTIAL.
FT DOMAIN 132 760 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 203 210 ATP (POTENTIAL).
FT METAL 425 425 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 426 426 BY SIMILARITY.
FT METAL 429 429 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 760 AA: 81985 MW: 81985 MW: AFPE8BFI0E7C6CF CRC64;

```

```

Query Match          5.6%; Score 127.5; DB 1; Length 760;
Best Local Similarity 22.6%; Pred. NO. 1.6;
Matches 88; Conservative 59; Mismatches 122; Indels 121; Gaps 23;

```

```

OY 64 RAVLAGPPATGKTALALGIAQELSGKVPFCMVGSE-----YSSSEVKTEYLMENF 116
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 197 KGVLYGPGTGTLLARAVAGEAG--VPFTISGSDFEVMEVGVGASRVR-----DLF 248
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 117 RRAIGIRIKENK--VYEEVTELSPEAEESTTGGYAKSISHVILSLKTVGKQLKIDS 174
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 249 EQA-----KQNSPCIIFFVEIDAIVAGRGAGIGGCHDER-----E 283
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 175 SIYDALIKKVAVD--VITYEANS-----GAVKRVGRCDSPATEYDLEAEYVPDIPK 225
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 284 QTLNQLVEMSGFGRAGVYILIAATNRPDILDPALLRGRD-----RQIPVSN 332
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 226 GE-----VHKR-KETIVDVTLDHDAANAPOG--GODIISLMO--MMKPRK--T 269
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 333 PDLAGRAVAVRHVSKGRPMADA--DLDGLAKRTVGTGADLANVINAEALILFARNGT 389
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 270 EIT-EKLRQELINKV-----NRYIDGIALVPGVLFIDEVHML-----DIECESYLN 316
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 390 VITGPALEEADVAVIGGRRRRRIISQOEKKI--TAYHEGSHTLAAMAMPDIE----- 440
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 317 RALESPISPIVILATNRGICNVRTDWTSPHGIPVDLDRLVIIIRIETTGPTMIQILAI 376
DB : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 441 -----PIYKVTILLNGR-----TGHRVAVPEEDKGLTRSEMI--AQLVFAMGC 483
OY 377 RAQVEEIDMDEESLAVIGRIGQOTSLRAHAI 406
DB 484 RA-AEELVREPTTGAVSDIEQATKIRASW 512

```

```

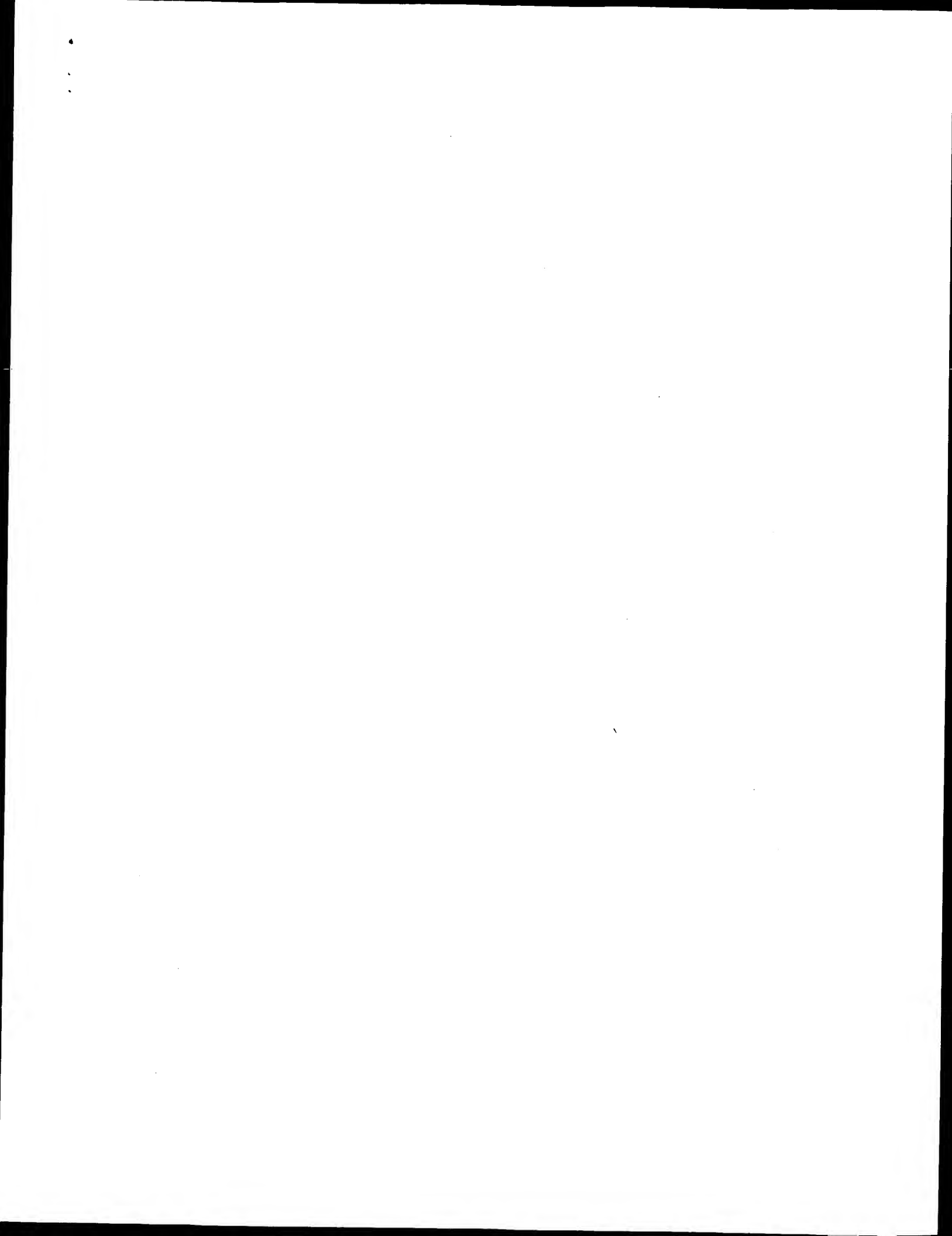
Search completed: November 13, 2001, 07:02:56
Job time: 342 sec

```

Wed Nov 14 08:34:18 2001

us-09-589-510-4.std.rsp

Page 11







[illegible]

REFERENCE 1 (bases 1 to 1427)  
 AUTHORS van Deursen, F.J., Shahi, S.K., Turner, C.M.R., Hartmann, C.,  
 Guerra-Giraldez, C., Matthews, K.R. and Clayton, C.E.  
 TITLE Characterisation of the growth and differentiation in vivo and in  
 vitro of bloodstream-form Trypanosoma brucei strain TREU 927  
 JOURNAL Mol. Biochem. Parasitol. 112 (2), 163-171 (2001)  
 COMMENT Contact: Shahi, S.K.  
 ZMBH (Centre for Molecular Biology)  
 Heidelberg University  
 Im Neuenheimer Feld 282, 69120 Heidelberg, Germany  
 POLYA-No.

FEATURES  
 source 1. 1427  
 /organism="Trypanosoma brucei"  
 /strain="927"  
 /db\_xref="taxon:5691"  
 /clone\_id="Trypanosoma brucei strain 927 ESTs"

BASE COUNT 387 a 322 c 424 g 294 t  
 ORIGIN

alignment scores:  
 quality: 1601.00 Length: 454  
 Ratio: 3.895 Gaps: 0  
 Percent Similarity: 90.529 Percent Identity: 67.841

alignment block:  
 US-09-589-510-4 x AJ131743 ..

Align seg 1/1 to: AJ131743 from: 1 to: 1427

```

1 MetArgIleGluIuValGlnSerThrSerLysLysGlnArgIleAlaThr 17
   ::::|||||:::|||||:::|||||:::|||||:::|||||:::
52 ATCAAAATTGAGGAGTTATTCACACCAAGAGAAAGAGAGCTGTACAGC 101
   ::::|||||:::|||||:::|||||:::|||||:::|||||:::
17 rHisThrHisIleLysGlyLysGlyLysLeuAspAlaAsnGlyMetAlaIleA 34
   :|||:::|||||:::|||||:::|||||:::|||||:::
102 TCACAGCAGCTGTGAAGGCGCTCGACTTGC. GCCGATGGCGGTAGCAAGC 150
   :|||:::|||||:::|||||:::|||||:::|||||:::
34 IaleuAlaIleAlaGlyPheValGlyGlnAlaIleAlaIleArgGluAlaIleGly 50
   ||| |||||:::|||||:::|||||:::|||||:::|||||:::
151 CGACTCGCAGATGATTGTTCGACAAAGTAAAGCCGTGAGCGCCGCTGGC 200
   ||| |||||:::|||||:::|||||:::|||||:::|||||:::
51 leuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaValIle 67
   ::::|||||:::|||||:::|||||:::|||||:::|||||:::
201 ATTGTGGTGGAGCTGACACGAAAGAAATGCGCGGTGCTGCTGCT 250
   ::::|||||:::|||||:::|||||:::|||||:::|||||:::
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaIleGlyIleA 84
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
251 CTTTGCAGGGCCCCCGGAGACCGGAGAAAGCGCACTGGCATTTGGGTGG 300
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
84 IagIleGluIleuGlySerLysValProPheCysPrometValGlySerGlu 100
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
301 CAAAGAACTCGGGGCCAAAGTTCATTCCTCCGATGTTGTAAGTAA 350
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
101 ValTyrSerSerGluValLysLysThrGluValIleuMetGluAsnPheAr 117
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
351 GTGTACAGCGCGGAGGTGAGAGAGGAGGTCTTCATGAGAAACTTCG 400
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
117 gArGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
401 CCGGGCGATTGGACTGCGATCAAGAGACGAGAGAGGTACAGAGGTG 450
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
134 IuValThrIleuLeuSerProGluIuValGluSerThrThrGlyGlyTyr 150
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
451 AGGTGACGAACTGCGTGCAGAGAAAGGATTAACCCGTTGAGAGGTAT 500
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
501 GGAAGTCATCTCCACGCTCATTTGTGACGCTCAATCGGTCAAGGCTC 550
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```

551 CAGACGTTGAAGCTTGATGCCCAATCTACGAAAGCTACAGAGAGAGA 600
184 ySValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
601 AAGTGGCGGTTGGTACGCTCATCTACATTGAGAGTACAGAGGCGGCTC 650
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluVal 217
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
651 AAGCGCTGGCGCGATCGGATGCTTATATTGGGAGCCAGCACTCGAGGC 700
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
217 agIuGluTyrValProIleProLysGlyLysValHisLysLysGlu 234
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
701 TGATGAGTATGTTCCGCTTCTTAAGGCGATGTCACAAAGAGAGAGA 750
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
234 IeValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
751 TTTATCAAGACGTAAACGCTTACAGATCTCGACGAGCCAAATGCTAAACA 800
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLeuLysProAr 267
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
801 AACCAAGGTCAGACGCACTTTCATTTGATGATGAGCTCATGAACAAAA 850
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
851 GAAACAGAAATTAACGAGAGAACTCCGCAATGAATAAACAAGTTGTGA 900
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
284 sNArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
901 ACAATATCATTTGACCAAGCGCTTCTGAGTTGGTCCCGCGCTTTTGTTC 950
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
951 ATTACAGAGGTCATATGTTAGACATAGAGTGTATTACATTCTTAATAVA 1000
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1001 GGCCCTGGAATCCACTGCTGCTCTGTTGTTATTTTTCACCAACAGGG 1050
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
334 LylIleCysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350
   ||| |||||:::|||||:::|||||:::|||||:::|||||:::
1051 GCAGTTGCGCATCAGAGGAGGAGGATGTGCGTCCCGCAGCGCATCCCG 1100
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPro 367
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1101 ACTGATCTTCTCGACCGGTTGCTGATGATGTGCGCACCGCAATTACAGCAT 1150
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
367 cThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGluIleA 384
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1151 TGAGAGGTGTGGCCATTGTGGACATTCGCCGCGAGTGAAGGTGCA 1200
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
384 sPmetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGlnGlnThr 400
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1201 GCGTATCGGATGCGGCTTTGGAGCTTCCGTTCAAAATCGCGAGCCGAGC 1250
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerTy 417
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1251 TCACCTTCGTTACGTGCGCAGCTGTGACACCAAGCGCTTATTTACGAGA 1300
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
417 sThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluValValSerG 434
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1301 AACCAACGGCGCATTCACCATTTGAAGTGGAGAGATGTACTTACTGTATG 1350
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
434 LysLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGluGln 450
   :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1351 GCCTTTTCAAGAGCAAAAGCCCTGCGGAGATGCTCCACAGAGAACCA 1400
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
451 GluArgTyrIle 454
   |||:::
1401 GAGGATTACGTC 1412
   ||:::
seq_name: gb_est29:AL553383
seq_documentation_block:

```

LOCUS AL553383 994 bp mRNA EST 16-FEB-2001  
 DEFINITION AL553383 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1075YM04 5  
 prime, mRNA sequence.  
 ACCESSION AL553383  
 VERSION AL553383.1 GI:12893166  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 994)  
 AUTHORS Li W.B., Gruber,C., Jessee,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 Source  
 1..994  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1075YM04"  
 /clone\_lib="LTI\_NFL006\_PL2"  
 /issue\_type="placenta"  
 /note="Vector: pCMVSPORT 6; site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : filiang@life-tech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 281 a 207 c 307 g 196 t 3 others  
 ORIGIN

alignment\_scores:  
 Quality: 1218.00 Length: 312  
 Ratio: 4.157 Gaps: 1  
 Percent Similarity: 93.910 Percent Identity: 78.526

alignment\_block:

US-09-589-510-4 x AL553383 ..

Align seg 1/1 to: AL553383 from: 1 to: 994

```

1 MetArgIleGluGluValGlnSerThrSerLysGlnArgTleAath 17
|||||
63 ATGAAGATTGAGGAGGTGAAGAGTACTACGAAAGCGCGCTCGCTC 112
|||||
17 rHisThrHisIleLysGlyLeuAspAlaAsnGlyMetAlaIleA 34
|||||
113 CCACAGCCACGTGAAGGCTGGGCTGACGAGAGCGGCTTGCCAAAC 162
|||||
34 LalaAlaIaIaGlyPheValGlyGlnAlaIaIaIaIaIaIaIaIa 50
|||||
163 AGGGGGCGCTCAGGCTGTGGGCGAGGAGACCGCGAGGCGCTATG 212
|||||
51 LeuAlaValAspMetIleArgGlnLysMetAlaGlyValAlaVal 67
|||||
213 GTCATAGTGAATTAATCAAAAGCAGAAATGCTGGAAGAGCTGCTT 262
|||||
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuAla 84
|||||
263 GTTGGCGGACCTCTGGAACCTGCAAGACACTGCTGCTGCTATG 312
|||||
84 LaglnLueGlySerLysValProPheCysProMetValGlySerGlu 100
|||||
313 CTCAGGAGCTGGTACTAGTCCCTTCTGCCAATGCTGGGAGTGA 362

```

```

101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
|||||
363 GTTACTCACTGAGTACGATCAGAAAGACAGAGGCTGCTGATGAGACTCCG 412
|||||
117 gArGAlaIleGlyLeuArgTleLysGlyLysGlyValLysGlyGlyG 134
|||||
413 CAGGGCCATTGGGCTGGGAATTAAGAGACCAAGGAAGTTTATGAAAGTG 462
|||||
134 LuValThrGluLeuSerProGluGluAlaGlySerThrThrGlyGlyTyr 150
|||||
463 AAGTCACAGAGTAACTCCGCTGTGACAGACAGAAATCCATGGAGAGATAT 512
|||||
151 AlalysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
|||||
513 GGCAAAACATTAGCCATGTGATCATAGACATCAAAACAGCCAAAGAAC 562
|||||
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
|||||
563 CAACAGATTGAAGTGGACCCAGCATTTTGTGAAGTTTGCAAGAAAGAGC 612
|||||
184 ysValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
|||||
613 GAATAGACCTGAGATGTGATTACATTGAAGCCAAAGTGGGCGCTG 662
|||||
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217
|||||
663 AAGAGCAGGCGAGGTGTGATCATCTATGCCACAGAAATTCGACCTGAAGC 712
|||||
217 agLugluTyrValProIleProLysGlyGluValHisLysLysGlyGlu 234
|||||
713 TCAGAGTATGTCCCTTGCCAAAGGGAGTGTGCACAAAGAAAGAA 762
|||||
234 lvaValGlnAspValThrLeuHisAspLeuAspAlaIaIaIaIaIaIa 250
|||||
763 TCATCTCAAGATGTGACCTTGATCATCTTGATGTGGCTAATGCCGCGGCC 812
|||||
251 Gln.clyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysPro 267
|||||
813 CAGGGGGGACAGATATCTCTGTCATGATGGCGCAGTATGAAGCCAA 862
|||||
267 rGlyThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValVal 283
|||||
863 AGAAGACAGAAATCACAGACAACTTCGAGGAGGAGATTAT.AAGGTGGTG 911
|||||
284 AsnArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPh 300
|||||
912 AACAGTACATGACCAAGGSAATGCTGAGCTGCTGCCGGTGTCTGTT 961
|||||
300 eLlAspGluValHisMetLeuAspIleGluCys 311
|||||
962 TGTGATGAGGTGTC.ATGCTGGACATGTGAGTGC 994
|||||
seq_name: gb_est29:AL542244
seq_documentation_block:
LOCUS AL542244 919 bp mRNA EST 16-FEB-2001
DEFINITION AL542244 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE08YGI2 5 prime
prime, mRNA sequence.
ACCESSION AL542244
VERSION AL542244.1 GI:12874097
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS Li W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

# FEATURES

## SOURCE

Location/Qualifiers  
 1..919  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE08YG12"  
 /clone\_1b="LTI-FL002\_PL1"  
 /lab\_host="DH10B"  
 /note="Organ: placenta; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 274 a 188 c 279 g 175 t 3 others  
 ORIGIN

alignment\_scores:  
 Quality: 1181.00 Length: 301  
 Ratio: 4.188 Gaps: 1  
 Percent Similarity: 93.688 Percent Identity: 77.741

alignment\_block:  
 US-09-589-510-4 x AL542244 ..

Align seg 1/1 to: AL542244 from: 1 to: 919

```

1 MetArlleGlulValAlnSerThrSerLysLysGlnArgIleAlaIrh 17
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 ATGAAGATTGAGAGGTGAGAGACACTACGAGACCCAGCATGCCCTC 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rHstHhIleIleLysGlyLeu.GlyLeuAspAlaAsnGlyMetAlaIle 33
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
70 CCACAGCCAGCTGAAGAGGCTGGAGACAGAGCCGCTGGGCCAG 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
34 AlaleuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaIle 50
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
120 CAGGCGGCTCAGGCGCTGTGGGCGCAGAGACGCGGAGAGCATGTG 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
50 yLeuAlaValAspMetIleArgGlnLysMetAlaGlyArgAlaValI 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
170 CGTCATAGTGAATTAATCAAAACCAAAATGGCTGGAGAGCTGTCT 219
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
67 eulLeuAlaGlyProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 83
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
220 TGTGGGAGGACCTCTCGAAGACTGGCAAGACGCTGGCTGGGCTATT 269
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
84 AlaglnGluLeuGlySerLysValProPheCysProMetValGlySerI 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
270 GCTCAGAGAGCTGGGTAGTAAGTCCCTCTGCCCAATGGTGGGAGTGA 319
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
100 uValLysSerSerGluValLysLysThrGluValLeuMetGluAsnPhe 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
320 AGTTTACTCACTGAKATCAAGAAAGACAGAGGTCTGATGAGAACTTC 369
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
117 rAArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGlyI 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
370 GGAGGCC.ATTGGGCTGGCAATTAAGAGCAAGAGAAATTAAGAGGT 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
134 GluValThrGluLeuSerProGluGluAlaGlnSerThrThrGlyGly 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GAAGTACACGAGCTTAAGTCCGTTGAGACAGAGATCCCATGGAGGATA 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
150 rAlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGly 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
469 TGGCAAAACCATTAAGCATGTGATGACATCAAAACCAAGCAAGGAA 518
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
167 hrlLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLys 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

519 CCAACAGTTGAACCTGAGCCCAACATTTTGAAGTTTGACAAAGAG 568
184 LysValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAla 200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
569 CCAGTAGAAGACCTGAGATGTGATTTACATGACAGCAAGTGGGGCGT 618
200 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGlu 217
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
619 GAAGAGCGCGGCGAGGTGTGATACCTATGCCACAGAAATTCACCTGA 668
217 lAglnGluTyrValProIleProLysGlyGluValHisLysLysGlu 233
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
669 CCGAAGATGATGTCCCTTGGCCAAAGGGAGATGTGCACAAAGAAAGAA 718
234 lIleValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGln 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
719 ATCATCAAGATGTGACCTTGATGATCTTGATGATGCGCTAATGCCGG 768
250 oGlnGlyGlyGlnAspIleLeuSerLeuMetLysGlnMetLysProA 267
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
769 CCAGGCGGCGACAAAGATATCTCTGATGATGCGCCAGCTAATGAAGCCA 818
267 rGlyThrGluIleThrGluLysLysArgGlnGluIleAsnLysValVal 283
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
819 AGAAGACAGAAATCACAGCAAACTTCAGGGGAGATTAATTAAGTGGTG 868
284 AsnArgTyrIleAspGlnGlyIleAlaGluLeuValProGlyValLeu 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
869 AACAGTACATGACAGCGGATTTCTGAGCTGTGTCCCGCTCTCTGT 918
300 e 300
919 t 919

```

seq\_name: gb\_est69:BE035216

seq\_documentation\_block:

LOCUS BE035216 1068 bp mRNA EST 07-JUN-2000  
 DEFINITION MO01E02 MO Mesembryanthemum crystallinum cDNA 5' similar to  
 ttp-interacting protein (tip49), mRNA sequence.

ACCESSION BE035216  
 VERSION BE035216.1 GI:8330340

KEYWORDS common ice plant.  
 EST.

SOURCE common ice plant.  
 ORGANISM Mesembryanthemum crystallinum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.

1 (bases 1 to 1068)  
 Bohnert,H.J., Bohnert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira,  
 H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,  
 Scara,G., Wheeler,M. and Zepeda,G.R.

Functional Genomics of Plant Stress Tolerance  
 Unpublished (2000)

CONTACT: Michalowski,C.B.  
 TITLE JOURNAL

COMMENT University of Arizona  
 Bio Sciences West room 513, Tucson, AZ 85721, USA  
 Tel: 520-621-7982  
 Fax: 520-621-1697  
 Email: cbm@u.arizona.edu

An open reading frame exists.  
 Insert Length: 1 Std Error: 0.00.

FEATURES  
 Location/Qualifiers

source 1..1068  
 /organism="Mesembryanthemum crystallinum"

/db\_xref="taxon:3544"  
 /clone\_1b="MO"

/tissue\_type="apical meristem and leaf primordia"  
 /dev\_stage="5 weeks"

/note="no stress"  
 BASE COUNT 314 a 224 c 281 g 248 t 1 others  
 ORIGIN

## alignment\_scores:

Quality: 1143.00 Length: 295  
 Ratio: 4.218 Gaps: 5  
 Percent Similarity: 91.864 Percent Identity: 84.746

## alignment\_block:

US-09-589-510-4 x BE035216 ..

Align seg 1/1 to: BE035216 from: 1 to: 1068

```

1 MetArgIleGluGluValGlnSerThrSerLysLysGlnArgIleAlaThr 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 ATGAAAGATTGAAGAAAGTTACAGTACACAAAGAACACAGCTATTGCCAC 242
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rHsThHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 TCACACCCACATCAAAAGCCCTTGCCCTCAGACGCCAACGGGAATGCCGTAC 292
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 lAlaValAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaGly 50
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 CCTGGCTGCTGGCTTTGTGGTCAGACAGACAGGAGGAGACCTGGA 342
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 lAlaValAlaSerMetIleArgGlnLysMetAlaGlyAlaValIle 67
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
343 CTGTACTTGTATGATGACGACAAAGAAATGGCTGGCTGACGACCTTCT 392
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 lAlaValAlaGlyProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
393 ACTTGGTGGCCCTCCCGGTACTGGCAGACAGCGCATTCGACCTTGGGATTG 442
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 lAlaGlnLeuGlySerLysValProPheCysProMetValGlySerGlu 100
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
443 CTCAGAGAGCTTGGAAAGCAAGGTTCCCTCTGCTATGTTGGCTCAGAA 492
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
493 GTTACTCATCAGAAAGTAAGAAAGCTAAGTTCTTATGAGAAATTTTAA 542
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 gATgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
543 ACAGAGATTGGTCTACGATCAAGAAACAAAGAGGTTATCAAGAGAG 592
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 lValThrGluLeuSerProGluGluAlaGluSerThrThyGlyTyr 150
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
593 AGGTGACAGAGCTCTCCCGAGAGAGAGGCTGAGAGTTCGACGTCGCTTAC 642
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 lAlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
643 ACTAGAGCATTTAGTCATGTAATTATTTGGTTAAAACTGTCAAAAGAGAC 692
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
693 CAAGAGAGCTGAATTTAGACCTCATATATAGATGCTTATTAAGGAA 742
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 ySValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
743 AGGTGCTGCTGGAGATGTTATTTAATTAAGGAGAACATGGGGCAGT 792
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGlu 216
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
793 AAAAAGGGTGGGAGAGTGGCTTTTGGACAGAAATTTGACTTTGAG 842
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 lAlaGluGluTyrValProIlePhePolysGlyGluValHisLysLys 232
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
843 GGTGAAGAAATATTGTCCACTTCCCAAGAGAGAGGCTTTTAAGAAAG 892
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 GlutIleValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAla 924
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
893 GAGAAATGTCAGATTAACCTACATGATCCGGATGCTGCAAGGCTTAA 942
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 lInProGlnGlyGlnAspIleLeuSerLeuMetGlyGlnKetMetLys 265
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

943 GGCTCAAGGTGGCCAGACATATTGCTTTATTGGGCCAGATGATGAG 992

266 ProArgLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLys 282

993 CCCACGAAACGAAATTCGGGTTAACTACCGCAGAGAAATTAACAAAGT 1041

282 lValAsnArgTyrIleAspGluGly 290

1042 AGTAAATTGGTTGTGATGAAGGC 1066

seq\_name: gb\_est29:AL557272

seq\_documentation\_block:

LOCUS AL557272 962 bp mRNA EST 16-FEB-2001

DEFINITION AL557272 LTI\_FL012\_Tc1 Homo sapiens cDNA clone CS0DH004YM17 5 prime

ACCESSION AL557272

VERSION AL557272.1 GI:12900714

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 962)

AUTHORS Li W.B., Gruber C., Jessee J. and Polyes D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

source

1..962

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DH004YM17"

/clone\_1lb="LTI\_FL012\_Tc1"

/tissue\_type="T cells from T cell leukemia"

/lab\_host="DH10B"

/note="vector: PCWSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-stranded cDNA was digested with Not I and cloned

into the Not I and Eco RV sites of the pCMVSPORT 6 vector.

Library was constructed by Life Technologies. Contact :

Feng Liang Life Technologies, a division of Invitrogen

9800 Medical Center Drive Rockville, Maryland 20850, USA

fax : (1) 301 610 8371 Email : lifang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 278 a 201 c 295 g 185 t 3 others

## alignment\_scores:

Quality: 1118.50 Length: 296  
 Ratio: 4.097 Gaps: 2  
 Percent Similarity: 92.230 Percent Identity: 76.351

alignment\_block:

US-09-589-510-4 x AL557272 ..

Align seg 1/1 to: AL557272 from: 1 to: 962

```

1 MetArgIleGluGluValGlnSerThrSerLysLysGlnArgIleAlaThr 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 ATGAAAGATTGAAGAGGTTGAGAGGCTACGAAAGACGACAGGATGCTTC 125
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rHsThHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 CCACAGCCACAGTGAAGAGGGGTGGGCTGACAGAGCGGCTTGGCCAAAG 175
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 lAlaValAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaGly 50
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 AGCGGCTTCAGGGCTTGGGCGCAGAGAACGCGCGAGAGGACATGTGGC 225
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

51 LeuAlaValAspMetIleArgGlnLysMetAlaGlyArgAlaValLeu 67
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
226 GTCAATAGATTAATTAATCAAGAGAAATAGCTGGAAGACTGTCTT 275
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
276 GTTGAGAGAGACCTCTGGAAGCTGGCAAGACAGCTGTGGCTGCTATTG 325
84 IagGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
326 CTCAGAGACTGGGTAGTAAGTCCCTTCTGCCCAATAGTGGGAGAGAA 375
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
376 GTTACTCACTACAGATCAAGAGACAGAGGTGTGTATGGAGAACTTCCG 425
117 gArGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
426 CAGGCC.ATTGGGCTGGCAATTAAGAGACCAAGAAAGTTATGAGGTG 474
134 IuValThrGluLeuSerProGluGluIleGluSerThrThrGlyTyr 150
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
475 AAGTCACAGAGCTAATCCGTGTGAGACAGAAATCCCATGGAGAGATAT 524
151 AlAlysSerLysSerHisValIleIleSerLeuLysThrValLysGlyTh 167
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
525 GGCAGAACCATTAAGCATCTGTATCTAGAGCTCAAAACAGCCAAAGAAC 574
167 rIysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
575 CAACAGCTTGAAGACAGACCCGACATTTTGAAGTTTGAGAAAGAGC 624
184 ysaValAlaValGlyAspValIleTyrIleGluAlaAsnSecGlyAlaVal 200
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
625 GACTAGAGAGTGAGATGTGATTACATTGAAGCCACAGAGGGGCGCTG 674
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
675 AAGAGGAGGCGAGTGTGTATCTGTACACAGAAATTCGACCTTGAAGC 724
217 agLugluTyrValProIleProLysGlyLysValHisLysLysGluI 234
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
725 TGAAGATGATGTCCTCTGCAGAAAGGGATGTGCACAAAAGAAAGAAA 774
234 IeValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnIleGlnPro 250
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
775 TCATCCAGATGTGACCTTGACATGTGATGTGCTAATGCGGCGCC 824
251 GlnGlyLysGlnAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
825 CAGGGGAGACAGATATCTGTCCATGATGCGACATATGAGGMAAA 874
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
875 GAAAGACGAATTCAGACAACTTCAGGGGAGATTAAATAGTGCTGA 924
284 snArg.TyrIleAspGluGlyIleAlaGluLeuVal 295
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
925 MAAGTACAT...CGACAAGGCATTGTGAGCTGTGC 957
seq_name: gb_est29:AL559327
seq_documentation_block: 893 bp mRNA EST 16-FEB-2001
LOCUS AL559327 LYL_NFL008_Tc2 Homo sapiens cDNA CS0D012Y013 5
DEFINITION prime, mRNA sequence.
ACCESSION AL559327
VERSION AL559327.1 GI:12904718
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE 1 (bases 1 to 893)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..893
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D012Y013"
/clone_lib="LTL_NFL008_Tc2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 265 a 186 c 269 g 172 t 1 others
ORIGIN

alignment_scores:
Quality: 1106.00 Length: 280
Ratio: 4.205 Gaps: 0
Percent Similarity: 93.929 Percent Identity: 77.500

alignment_block:
us-09-589-510-4 x AL559327
Align seg 1/1 to: AL559327 from: 1 to: 893

1 MetArgIleGluGluValGlnSerThrSerLysGlnArgIleAlaTh 17
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
53 ATGAAGATTGAGAGAGGTGAAGAGACACTGAGAGACGCGATGCGCTC 102
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
103 CCACAGCCAGGTGAAGGCGCTGGGCGTGGACGAGAGCGCTTGCCCAAGC 152
34 IeLeuAlaIleGlyPheValGlyGlnAlaAlaIleArgGluAlaIleGly 50
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
153 AGCGGCGCTCAGGCGCTTGTGGCCAGAGACGCGCGAGAGGCATGTGGC 202
51 LeuAlaValAspMetIleArgGlnLysMetAlaGlyArgAlaValLeu 67
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
203 GTCAATAGATTAATTAATCAAGAGAAATAGCTGGAAGAGCTGTCTT 252
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
253 GTTGAGAGAGACCTCTGGAAGCTGGCAAGACAGCTGTGGCTGCTATTG 302
84 IagGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
303 CTCAGAGACTGGGTAGTAAGTCCCTTCTGCCCAATAGTGGGAGAGTAA 352
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
353 GTTACTCACTACAGATCAAGAGACAGAGGTGTGTATGGAGAACTTCCG 402
117 gArGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
403 CAGGCC.ATTGGGCTGGCAATTAAGAGACCAAGAAAGTTATGAAGGTG 451

```



134 luValThrGluLeuSerProGluGluAgluSerThrThrGlyGlyTyr 150  
 452 AAGTCACAGAGCTAACCTCGTGTGACAGCAAGATCCATCGAGGACATAT 501  
 151 AlalysSerIleSerHisValIleIleSerLeuLysThrValGlyGly 167  
 502 GGCMAAACCATTTAGCCATGTGATCATAGACTCAAAACAGCAAGAAC 551  
 167 rlyGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184  
 552 CAACAGATTGAACTGACCCCGCATTTTGAAGTTGCGAAGAAC 601  
 184 ySValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200  
 602 GATTAAGAGCTGGAGATGTGATTACATTGAAGCCACACAGTGGGCGCTG 651  
 201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
 652 AAGAGCGACAGGCGAGTGTGATACCTATGCCACAGAAATCGACCTTGAGAC 701  
 217 aglGluTyrValProIleProLysGlyGluValHisLysLysLysGlu 234  
 702 TGAAGAGTATGTCCCTTGCCAAAAGGGGAGTGTGCACAAAAGAAAGAA 751  
 234 lValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250  
 752 TCATCCAAATGTGACCTTGACCTTGATGATGTGGCTAATGCGCGGCC 801  
 251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyLysMetLysProArg 267  
 802 CAGGGGGGCAAGATATCTGTCATGATGGCCACGCAATGAAGCCAAA 851  
 267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsn 280  
 852 GAAGCAAGAAATCACAGCAAACTTCAGAGGGAGATTAAAT 891  
 seq\_name: gb\_est29:AL535830

seq\_documentation\_block: 877 bp mRNA EST 13-FEB-2001  
 LOCUS AL535830 LTI\_FL013\_Fbrn1 Homo sapiens cDNA clone CS0DF013Y110 5  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL535830  
 VERSION AL535830.1 GI:1279323  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 877)  
 AUTHORS Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES  
 source  
 1. 877  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF013Y110"  
 /dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
 /lab\_host="DH108"  
 /note="Organ: Fetal brain. Vector: pCMWSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMWSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville

BASE COUNT 258 a 185 c 262 g 166 t 6 others  
 ORIGIN

alignment\_scores:  
 Quality: 1100.00 Length: 276  
 Ratio: 4.264 Gaps: 0  
 Percent Similarity: 93.478 Percent Identity: 76.449

alignment\_block:  
 US-09-589-510-4 x AL535830

Align seg 1/1 to: AL535830 from: 1 to: 877

1 MetArgIleGluGluValGlnSerThrSerLysGlnArgIleAlaThr 17  
 44 ATGAGATGTGAGAGGTGAGAGCACTACGAGACGCGCATCGCTC 93  
 17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIle 34  
 94 CCACAGCCACGTGAAGGGGCTGGGCTGACGAGAGCGGCTGGCCAGC 143  
 34 lAlaValAlaGlyPheValGlyGlnAlaAlaAlaArgIleAlaGly 50  
 144 AGCGCGCTCCAGGCTGTGTGGCCAGAGACGCGCAGAGGATGTGGC 193  
 51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaVal 67  
 194 GTCATAGTAGAATTAATCAAAAGCAAGAAATGCTGGAAGATCTGCTT 243  
 67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84  
 244 GTTGGCAGACCTCTGGAAGCTGCAAGACAGCTGCTGCTGCTATTTG 293  
 84 lAglGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100  
 294 CTCAGAGAGCTGGTANTAAAGTCCCTTCTGCCCAATGTGGGAGTGA 343  
 101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheArg 117  
 344 GTTACTCACTCACTGATCAAGAGACAGAGGTCTGATGAGACTCTCG 393  
 117 gArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134  
 394 CAGGGCCATTGGCTGCAATTAAGACCAAGAAAGATTATGAAGGTG 443  
 134 luValThrGluLeuSerProGluGluAgluSerThrThrGlyGlyTyr 150  
 444 AAGTCACAGAGCTAACCTCGTGTGACAGCAAGATCCATCGAGGACATAT 493  
 151 AlalysSerIleSerHisValIleIleSerLeuLysThrValGlyGly 167  
 494 GGCMAAACCATTTAGCCATGTGATCATAGACTCAAAACAGCAAGAAC 543  
 167 rlyGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184  
 544 CAACAGATTGAACTGACCCCGCATTTTGAAGTTGCGAAGAAC 593  
 184 ySValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200  
 594 GATTAAGAGCTGGAGATGTGATTACATTGAAGCCACACAGTGGGCGCTG 643  
 201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
 644 AAGAGCGACAGGCGAGTGTGATACCTATGCCACAGAAATCGACCTTGAGAC 693  
 217 aglGluTyrValProIleProLysGlyGluValHisLysLysLysGlu 234  
 694 TGAAGAGTATNTCCCTTGCCAAAAGGGGAGTGTGCACAAAAGAAAGAA 743



234 leValGlnaspValThrLeuHisaspLeuaspAlaalaaspAlaGlnPro 250  
|||||  
744 CATTCGAAGATGTGACCTTGATGACCTTGATGCTGATGAGCGGCC 793  
251 GlnGlyGlnaspIleuSerIleuSerMetGlyGlnMetMetLysProAr 267  
|||||  
794 CAGGGGGGACAGATATCTGCTGATGATGCGCAGCTAATGAGCCAAA 843  
267 gLysThrGlnIleuThrGlnLysLeuArg 276  
|||||  
844 GAGACAGAAATCCACAGACAACTTCA 871  
seq\_name: gb\_est197.BG126632  
seq\_documentation\_block: 764 bp mRNA EST 31-JAN-2001  
LOCUS BG126632  
DEFINITION EST472278 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
CTOF12N8 5' sequence, mRNA sequence.  
ACCESSION BG126632  
VERSION BG126632.1 GI:12626820  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; eunasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 764)  
AUTHORS Van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,  
Hansen, C., Roming, C. and Tanksley, S.  
TITLE Generation of ESTs from tomato shoot/meristem tissue  
JOURNAL Unpublished (2001)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
FEATURES  
source  
1..764  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CTOF12N8"  
/clone\_lib="tomato shoot/meristem"  
/tissue\_type="shoot/meristem"  
/dev\_stage="developing shoots from 4-6wks old plants"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; Small expanding leaves from the growing tip were  
taken from greenhouse plants (4-6wks old TA496). Tissue  
was immediately frozen in liquid nitrogen."  
BASE COUNT 243 a 124 c 210 g 187 t  
ORIGIN  
alignment\_scores:  
Quality: 1077.00 Length: 246  
Ratio: 4.525 Gaps: 0  
Percent Similarity: 96.748 Percent Identity: 88.211  
alignment\_block:  
US-09-589-510-4 x BG126632 ..  
Align seq 1/1 to: BG126632 from: 1 to: 764  
1 MetArgIleGlnGluValGlnSerThrSerLysLysGlnArgIleAlaTh 17  
|||||  
22 ATGAAGATAGAGAGGTACATCAATCAATCAATCAATCAATGCTTAC 71  
|||||  
17 rHisThrHisIleLysGlyLeuGlyLeuaspAlaasnGlyMetAlaIleA 34  
|||||  
72 TCATATCATCATTAAGAGCCTTGCTTGTGAGCCAAATGAGAGGACACTTC 121

34 laLeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaGly 50  
|||||  
122 CATTCGACACTGGGTTTGTAGCTCAGGCGCAGCCAGAGAACTCGGGG 171  
51 leuAlaValaspMetIleArgGlnLysLysMetAlaGlyArgAlaValle 67  
|||||  
172 CTGTGGTGGATATGATACCGCAAAAGAAATGCTGGCGGCTTTACT 221  
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84  
|||||  
222 CCTAGCTGTCCTCCCTGCTACGGCGCAGACACTCTGCTCTTGGTATAT 271  
84 laGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100  
|||||  
272 CACAAGACCTTGGAAGCAGAGTTCCATTTGTCCATGTTGGTCCGGA 321  
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluaspPheAr 117  
|||||  
322 GTGTATTCATCAGAAAGTGAAGAAACTGAGCTTAATGAAACCTTCG 371  
117 gArGAlaIleGlyIleuArgIleLysGlnLysGluValTyrGluGlyG 134  
|||||  
372 GCGAGCTATTGCTCTCGTATGAGGAAATAGAGAGTTTATGAGAGAG 421  
134 luValThrGluLeuSerProGlnGluAlaGluSerThrThrGlyGlyTyr 150  
|||||  
422 AGGTGACTGACACTCTCCAGAAAGCGTACAGAGTGTGACAGTGGATAT 471  
151 AlAlaSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167  
|||||  
472 GGTAAAGCATTAAGCATGCTTAATGCGTTAAACACTGCAAGAGTAC 521  
167 rLysGlnLeuLysLeuaspSerSerIleTyraspAlaLeuLysGluL 184  
|||||  
522 CAACACGTTGAAGCTTGACCCACGATATATGATGCCATATTAAGAGA 571  
184 ysValAlaValGlyaspValIleTyrIleGluAlaaspSerGlyAlaVal 200  
|||||  
572 AGTACCTGCTGGTGTGATGATCATTTACATGATCAACAGTGGACAGAT 621  
201 LysAlaGlyAlaArgCysaspSerPheAlaThrGluTyrAspLeuGluAl 217  
|||||  
622 AAAAGAGTGGCAGAGGATGCTTTGGCACAGAAATTTGATCTTGAGCA 670  
217 agLugluTyrValProIleProLysGlyGluValHisLysLysGluI 234  
|||||  
671 .GAGAGTATGTTCCACTTCTTAAGAGAGGTTCACAGAGAAAGAGAA 720  
234 leValGlnaspValThrLeuHisaspLeuaspAlaAla 246  
|||||  
721 TAGTCCAGATGTTACATTCATGATGATCTTGATGCTGCA 758  
seq\_name: gb\_est28:AL514734  
seq\_documentation\_block:  
LOCUS AL514734 899 bp mRNA EST 13-FEB-2001  
DEFINITION AL514734 LTL.NFL006.PL2 Homo sapiens cDNA clone CL08B013ZH07 5  
prime, mRNA sequence.  
ACCESSION AL514734  
VERSION AL514734.1 GI:12778227  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 899)  
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).



## alignment\_scores:

Quality: 1056.00 Length: 285  
 Ratio: 4.093 Gaps: 1  
 Percent similarity: 90.526 Percent identity: 74.035

alignment\_block:  
 US-09-589-510-4 x BG397396 ..

Align seg 1/1 to: BG397396 from: 1 to: 899

```

93 PheCysPrometValGlySerGluValTyrSerSerGluValIleLysThr 109
|||||
3 TTCTGCCCAATGCTGGGAGTGAAGTTTACTTCACCTGAGATCAACAGAC 52
|||||
109 rGluValLeuMetGluAsnPheArgArgAlaIleGlyLeuArgIleLysG 126
|||||
53 AGAGGTGCTGATGAGAACTCCGAGGCGCATTTGGCTGGGAATAAAG 102
|||||
126 LuAsnLysGluValTyrGluGluValIleThrGluLeuSerProGluGlu 142
|||||
103 AGACCAAGAAAGTTATGAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 152
|||||
143 AlaGluSerThrThrGlyTyrAlaLysSerIleSerHisValIleIle 159
|||||
153 ACAGAGATTCCTCCATGGGAGATATGCGCAAAACCATTCACCTGATCAT 202
|||||
159 eSerLeuLysThrValLysGlyThrLysGluLeuLysLysSerSer 176
|||||
203 AGGACTCAAAACAGCCAAAGCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 252
|||||
176 LeTyrAspAlaLeuIleLysGluLysValAlaValAlaGlyAspValIleLys 192
|||||
253 TTTTGAAG..TTTGAAGAAAGAGCGATGAAGTGAAGTGAAGTGAAGTGAAGT 301
|||||
193 IleGluAlaAsnSerGlyAlaValLysArgValGlyArgCysAspSerPh 209
|||||
302 ATTGAAGCCACACAGTGGGCGCTGAGAGGCGAGGCGAGTGTGATACCTA 351
|||||
209 eAlaThrGluTyrAspLeuGluAlaGluGluLysValProIleProLysG 226
|||||
352 TCCACAGAAATTCGACCTTGAAGCTGAAGAGATGCTCCCTTGCACAAAG 401
|||||
226 LysGluValHisLysLysLysGluIleValGlnAspValThrLeuHisAsp 242
|||||
402 GGGATGTCACAAAAAAGAAAGAAATCATCAAGATGTGACCTTGCAATGAC 451
|||||
243 LeuAspAlaAlaAsnAlaGlnProGlnGlyGlnAspIleLeuSerLe 259
|||||
452 TTGGATGTGGCTAATGCCGCTGCCAGGGGGGACAAAGATATCTGTCAT 501
|||||
259 uMetGlyLysMetLysProArgLysThrGluIleThrGluLysLeu 276
|||||
502 GATGGCCAGCTAATGAAGCCAAAGAACAGAAATCAACAGCAAACTTC 551
|||||
276 rGlnGluIleAsnLysValValAsn..ArgTyrIleAspGluGlyIleAl 292
|||||
552 GAGGGAGATTAAATGAAGTGTGAACCAAGTACATCGACCGGGCATTTGC 601
|||||
292 aGluLeuValProGlyValLeuPheIleAspGluValHisMetLeuAspI 309
|||||
602 TGAGTGTGCTCCGGGTGCTGCTTTTGTGATGAGTCCACATGCTGCACA 651
|||||
309 LeGluCysPheSerTyrLeuAsnArgAlaLeuGluSerProLeuSerPro 325
|||||
652 TTGAAGTCTTCACTACCTGCACCGCGCTGAGATCTTTATCGCTCC 701
|||||
326 IleValIleLeuAlaThrAsnArgGlyIleCysAsnValArgGlyThrAs 342
|||||
702 ATTCGATCTTTGATCCACGAGGACCTTGTGATCAGGAGGACCTGAG 751
|||||
342 pMetThrSerProHisGlyIleProValAspLeuAspArgLeuValI 359

```

seq\_name: gb\_est29:AL551666

seq\_documentation\_block:

LOCUS AL551666 866 bp mRNA EST 16-FEB-2001

DEFINITION AL551666 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1063J05 5

prime, mRNA sequence.

ACCESSION AL551666

VERSION AL551666.1 GI:12889833

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 866)

Li W.B., Gruber C., Jesse J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished (2001)

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

location/Qualifiers

1..866

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0D1063J05"

/clone\_lib="LTI\_NFL006\_PL2"

/rissue-type="placenta"

/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA

was primed with a NotI-Oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 247 a 183 c 266 g 168 t 2 others

ORIGIN

alignment\_scores:

Quality: 1013.00 Length: 263

Ratio: 4.118 Gaps: 1

Percent similarity: 93.536 Percent identity: 76.806

alignment\_block:

US-09-589-510-4 x AL551666 ..

Align seg 1/1 to: AL551666 from: 1 to: 866

```

1 MetArgIleGluGluValAlaGlnSerThrSerLysLysGlnArgIleAlaThr 17
|||||
74 ATCAAGATTGAGAGAGTGAAGACACTACGAAAGACCAAGCGCATCCGCTC 123
|||||
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
|||||
124 CCACAGCCACGTGAAGAAGGCTGGGCTGGACGAGACGCGCTGGCCAAAC 173
|||||
34 laLeuAlaIleGlyPheValGlyGlnAlaAlaIleArgGlnAlaIleGly 50
|||||

```

```

174 AGCGGGCTCAGGCTTGTGGCCAGAGAACGCCGAGAGGCGATGTGGC 223
51 LeuAlaValAspMetIleArgIleLysMetAlaGlyArgAlaValIle 67
::: |||:::|||||:::|||||:::|||||:::|||||:::|||||
224 GTCATAGTGAATTAATCAAAAGCAAGAAATGGCGGAAGAGCTGTCTT 273
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
274 GTTGGCAGAGCTCTCGAAGCTGGCAAGACAGCTGTGGCTGTGCTATTG 323
84 IaGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
324 CTCAGAGAGCTGGTAGTAAGTCCCTCTGCCCAATGGGGAGACTGAA 373
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPhe 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
374 GTTTACTCAACTGAGATCAAGAGACAGAGTGTCTATGGAGACTTCGG 423
117 gATGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
424 CAGGGCCATTGGGCTCGAATTAAGAGACCAAGAAATTATGAAGTG 473
134 LuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyGly 150
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
474 AAGTCACAGAGCTAAGCTCCGTGAGACAGAGAAATCCCATAGGAGGATA 523
150 rAlaLysSerIleSerHisValIleIleSerLeuLysThrValGlyGly 167
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
524 TGGCAAAACCATTAAGCATGATGATGATGATGATGATGATGATGATGAT 573
167 hrLysGlnLeuLysLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 183
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
574 CCAACAGCTGAAGCTGACCCAGCCAGCATTTTGAAGTTTGGCAAGAG 623
184 LysValAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAla 200
::: |||::: |||::: |||::: |||::: |||::: |||::: |||
624 CGAGTGAAGAGCTGAGATGTGATTTACATGGAAGCCACAGTGGGGCCG 673
200 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGlu 217
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
674 GAGAGGCGACGACGCTGATACCTATGCGACAGAAATTCGACCTTGAG 723
217 IaGlnGluTyrValProIleProLysGlyLysValHisLysLysGlu 233
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
724 CTGAGAGGTATGTCCCTGCCAAAGGGGAGTGCACAAAGAAAGAA 773
234 IleValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnIle 250
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
774 ATCATCCAGAGATGTGACCTTGATGACTTGATGATGATGATGATGATG 823
250 oGlnGlyGlnAspIleLeuSerLeuMetGlyGln 262
| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
824 CAGGGGAGCAAGATATCTCTGTCATGATGAGGCGAG 859

```

seq\_name: gb\_est80:BE901271

seq documentation block: 939 bp mRNA EST 29-SEP-2000

LOCUS BE901271 939 bp mRNA EST 29-SEP-2000

DEFINITION 601675912F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3958683 5',

ACCESSION BE901271

VERSION BE901271.1 GI:10390287

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 939)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: [image.llnl.gov](http://image.llnl.gov)

Plate: LICM38 row: 9 column: 04

High quality sequence start: 4

High quality sequence stop: 750.

FEATURES

Location/Qualifiers

1..939

source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3958683"

/clone\_lib="NIH\_MGC\_21"

/tissue\_type="choriocarcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: Placenta; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 259 a 208 c 302 g 170 t

ORIGIN

alignment\_scores:

Quality: 1006.50 Length: 266

Ratio: 4.075 Gaps: 2

Percent Similarity: 92.857 Percent Identity: 74.436

alignment\_block:

US-09-589-510-4 x BE901271 ..

Align seg 1/1 to: BE901271 from: 1 to: 939

```

1 MetArgIleGlnGluValGlnSerThrSerLysGlnArgIleAlaThr 17
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
22 ATGAGAGATTGAGAGGATGAAGGACCTAGCAAGACGCGCATGCGCTC 71
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIle 34
::: |||::: |||::: |||::: |||::: |||::: |||::: |||
72 CCACAGCCACGTGAAGAGGCTGGGCTGGAGAGAGCGCTTGGCCAAAGC 121
34 IaLeuAlaAlaGlyPheValGlyGlnAlaAlaIleArgGluAlaGly 50
|||::: |||::: |||::: |||::: |||::: |||::: |||
122 AGGCGCGCTCAGGCGCTGTGGGCCAGAGAACGCGCGAGGCGATGTGCG 171
51 LeuAlaValAspMetIleArgIleLysLysMetAlaGlyArgAlaValIle 67
::: |||::: |||::: |||::: |||::: |||::: |||::: |||
172 GTCATAGTGAATTAATCAAAAGCAAGAAATGGCTGGAAGAGCTGTCTT 221
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
222 GTTGGCAGAGACTCTCTGGAAGCTGGCAAGACAGCTGTGCTGTGCTATTG 271
84 IaGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
272 CTCAGAGAGCTGGGTAGTAAGTCCCTCTGCCCAATGGTGGGAGAGTGA 321
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPhe 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
322 GTTTACTCAACTGAGATCAAGAGACAGAGTGTGCTGATGGAGAACTTCG 371
117 gATGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
372 CAGGGCCATTGGGCTCGAATTAAGAGACCAAGAAATTATGAAGTG 421
134 LuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyGly 150
|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```



```

701 ATCCAGATGTGACTTGACCTGATGTGGCTAATGCGGGGCCCA 750
251 ngllyglnaspilleuSerleuMetglglnMetMetlyspArglu 268
||||| :|||:|||||:|||||:|||||:|||||:
751 GGGGACACAGA...TATCTGGCCATGAAGGGGACAGTAAAGCAAGGA 796
268 ysthrglu 270
|||||
797 AGACCGCA 804

seq_name: gb_est71:BE204244

seq_documentation_block:
LOCUS BE204244 632 bp mRNA EST 05-SEP-2000
DEFINITION EST396920 KVO Medicago truncatula cDNA clone pkV0-14N12, mRNA
sequence.
ACCESSION BE204244
VERSION BE204244.1 GI:8747529
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 632)
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,
Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E., and
Fraser, C.M.
ESTs from uninoculated seedling roots of Medicago truncatula
Unpublished (1999)
Contact: VandenBosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: katemail.bio.tamu.edu
Texas A&M University name: T264016e
TIGR sequence name: MTGAP78TK
More information is available at:
http://chrysis.tamu.edu/medicago.
Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
FEATURES
source
1.632
Location/Qualifiers
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pkV0-14N12"
/clone_1b="KVO"
/tissue_type="Seedling roots"
/dev_stage="Immediately prior to inoculation with
Sinorhizobium meliloti (0 hour)"
/lab_host="E.coli strain XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-zap phage using Ex-assist
helper phage and propagated in XLOLR cells."

BASE COUNT 183 a 116 c 160 g 173 t
ORIGIN

alignment_scores:
Quality: 995.00 Length: 210
Ratio: 4.807 Gaps: 0
Percent Similarity: 98.571 Percent Identity: 92.381

alignment_block:
US-09-589-510-4 x BE204244 ..
Align seg 1/1 to: BE204244 from: 1 to: 632

```

```

170 leuLysleuaspSerSerIleTyrAspAlaLeuIleLysGluLysValAl 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 CTAAGTGGACCCACAGATATATGATGCTTGCATCAACAAAAAGAGAC 50
186 aValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaValLysArgV 203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 TCTCGGGAGTATATATATATGAGCAATTAAGTGGCGCTGTGAAAAAGG 100
203 aLgLyArgCysAspSerPheAlaIleThrGluTyrAspLeuGluAlaGlu 219
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 TTGGCCGAAGTGAATGCTTTGCTACTGAGTGTGACCTTGAAAGCTGAGAG 150
220 TyrValProIleProLysGlyIleValIleLysLysLysGluIleValGI 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 TATGTTCCACTTCCCAAGGAGAGGTTCCACACAAAAAAGAGTTGTCA 200
236 nasrValIThrLeuIleAspLeuAspAlaAlaAsnAlaGlnProGlnGly 253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 GATGTAAACCTACATGATCTGGATGCTGCCAACGGCGCACTCAAGGGG 250
253 LysIleAspIleLeuSerleuMetGlyGlnMetMetlyspProArgLysThr 269
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 GGCAGAGTATTTGTCTCTTATGGGCCAGATGATGAACCTAGCAAGACG 300
270 GluIleThrGluLysLeuArgGlnGluIleAsnLysValValAsnArgTy 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GAAATCAGTACAGAGTTGACACAAAGAAATTAATTAAGTTGTCAACGATA 350
286 rIleAspGluGlyIleAlaGluLeuValProGlyValLeuPheIleAspG 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 TATCGATGAAGGAGTGTGCAGAACCTGTCTCGGGGTTCTTTATTATGACG 400
303 IuValHisMetLeuAspIleGluCysPheSerTyrIleuAsnArgAlaLeu 319
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 AGGTGCATATGCTAGATATGAGAGTGTTCCTATCTTAATCTGCTTTA 450
320 GluSerProLeuSerProIleValIleLeuAlaThrAsnArgGlyIleCy 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 GAGAGCTCGTGTCTCCCAATAGTAATCTTCGCCCAAAATGAGAGATATG 500
336 sAsnValArgGlyThrAspMetThrSerProHisGlyIleProValAspL 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 CACCGTAAGAGGAGCTGATATGACACAGTCTCTCATGGCATACCTGTTGACC 550
353 euLeuAspArgLeuValIleIleArgThrGluThrTyrGlyProThrGlu 369
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 TATTGATCGATTGGTGAATTCGAAACACAAACCTAAGGCTCAAGCTGAA 600
370 MetIleGlnIleLeuAlaIleArgAlaGln 379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 ATATATCAGATTCTAGCTATCCGAGCTCAA 630

```

Wed Nov 14 08:34:19 2001

us-09-589-510-4.std.rst

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 07:57:39 ; Search time 2534.15 Seconds  
(without alignments)  
11670.318 Million cell updates/sec

Title: US-09-589-510-3  
1912  
Perfect score: 1 acccagcgctcgcaattt.....aaaaaaaaaaaaaaaa 1912  
Sequence:

Scoring table: OLIGO\_MUC  
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:\*

- 1: gb\_da1:\*
- 2: gb\_da2:\*
- 3: gb\_da3:\*
- 4: gb\_in1:\*
- 5: gb\_in2:\*
- 6: gb\_in3:\*
- 7: gb\_om:\*
- 8: gb\_ov:\*
- 9: gb\_pat1:\*
- 10: gb\_pat2:\*
- 11: gb\_ph:\*
- 12: gb\_pl1:\*
- 13: gb\_pl2:\*
- 14: gb\_pl3:\*
- 15: gb\_pl4:\*
- 16: em\_da1:\*
- 17: em\_da2:\*
- 18: em\_fun:\*
- 19: em\_htgo\_hum:\*
- 20: em\_htgo\_inv:\*
- 21: em\_htgo\_rod:\*
- 22: em\_htg\_hum1:\*
- 23: em\_htg\_hum2:\*
- 24: em\_htg\_hum3:\*
- 25: em\_htg\_hum4:\*
- 26: em\_htg\_hum5:\*
- 27: em\_htg\_hum6:\*
- 28: em\_htg\_hum7:\*
- 29: em\_htg\_hum8:\*
- 30: em\_htg\_inv1:\*
- 31: em\_htg\_inv2:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_rod:\*
- 34: em\_hum1:\*
- 35: em\_hum2:\*
- 36: em\_hum3:\*
- 37: em\_hum4:\*
- 38: em\_hum5:\*
- 39: em\_hum6:\*
- 40: em\_hum7:\*
- 41: em\_in:\*
- 42: em\_om:\*
- 43: em\_or:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_pl:\*

48: em\_ro:\*

49: em\_sts:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_v1:\*

53: gb\_sts1:\*

54: gb\_sts2:\*

55: gb\_sts3:\*

56: gb\_sy:\*

57: gb\_un:\*

58: gb\_v11:\*

59: gb\_v12:\*

60: gb\_htg1:\*

61: gb\_htg2:\*

62: gb\_htg3:\*

63: gb\_htg4:\*

64: gb\_htg5:\*

65: gb\_htg6:\*

66: gb\_htg7:\*

67: gb\_htg8:\*

68: gb\_htg9:\*

69: gb\_htg10:\*

70: gb\_htg11:\*

71: gb\_htg12:\*

72: gb\_htg13:\*

73: gb\_htg14:\*

74: gb\_htg15:\*

75: gb\_htg16:\*

76: gb\_htg17:\*

77: gb\_htg18:\*

78: gb\_htg19:\*

79: gb\_htg20:\*

80: gb\_htg21:\*

81: gb\_htg22:\*

82: gb\_htg23:\*

83: gb\_htg24:\*

84: gb\_htg25:\*

85: gb\_pr1:\*

86: gb\_pr2:\*

87: gb\_pr3:\*

88: gb\_pr4:\*

89: gb\_pr5:\*

90: gb\_pr6:\*

91: gb\_pr7:\*

92: gb\_pr8:\*

93: gb\_pr9:\*

94: gb\_r01:\*

95: gb\_r02:\*

96: gb\_in4:\*

97: gb\_pr10:\*

98: em\_da3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1912	100.0	1912	10	AX077226 Sequence
2	1080	56.5	1869	10	AX077232 Sequence
3	576	30.1	1845	10	AX077224 Sequence
4	94	4.9	1886	10	AX077228 Sequence
5	94	4.9	1898	10	AX077230 Sequence
6	2.1	171809	83	AP003271	AP003271 Oryza sat
7	35	1.8	1173	53	CNS06LVV
8	34	1.8	55	10	129928 Sequence 41

C	18	34	1.8	71	97	S76508	
C	19	34	1.8	72	97	S76510	
C	20	34	1.8	73	97	A08915	H.sapiens (
C	21	34	1.8	77	9	A08908	H.sapiens (
C	22	34	1.8	79	10	AX099452	H.sapiens (
C	23	34	1.8	83	10	A08907	H.sapiens (
C	24	34	1.8	88	10	AX098165	H.sapiens (
C	25	34	1.8	91	6	DD1ACTINBA	M25215 Dictyostell
C	26	34	1.8	96	9	A08909	H.sapiens (
C	27	34	1.8	100	95	RNU12531	U12531 Rattus norv
C	28	34	1.8	102	9	A08911	A08911 H.sapiens (
C	29	34	1.8	104	9	A08910	A08910 H.sapiens (
C	30	34	1.8	107	9	A08912	A08912 H.sapiens (
C	31	34	1.8	114	10	122013	AR014063 Sequence
C	32	34	1.8	120	10	AR050959	122013 Sequence 99
C	33	34	1.8	120	10	151760	AR050955 Sequence
C	34	34	1.8	120	10	184386	I51760 Sequence 28
C	35	34	1.8	127	97	IROESTT43	I84386 Sequence 44
C	36	34	1.8	132	9	A08913	AL839935 Homo sapi
C	37	34	1.8	140	6	DMRRL	A08913 H.sapiens (
C	38	34	1.8	140	10	I48978	X08015 Drosophila
C	39	34	1.8	144	10	I89947	I48978 Sequence 5
C	40	34	1.8	161	9	A08916	I89947 Sequence 26
C	41	34	1.8	196	97	IROEST145	A08916 H.sapiens (
C	42	34	1.8	200	54	G38015	AL839939 Homo sapi
C	43	34	1.8	200	94	AF168594	G38015 C1M31 Plasm
C	44	34	1.8	210	7	AF250317	AF168594 Rattus no
C	45	34	1.8	216	10	I80062	AF250317 Canis fam
C	46	34	1.8	224	92	HS153811R	I80062 Sequence 34
C	47	34	1.8	225	8	S7878952	S56913 H.sapiens C
C	48	34	1.8	227	85	AB0112260	S78786 cGATA-3 (ch
C	49	34	1.8	229	9	AX062568	AB0112260 Homo sapi
C	50	34	1.8	236	54	G37836	AX062568 Sequence
C	51	34	1.8	240	97	I48979	G37836 wh1sapi Pla
C	52	34	1.8	247	6	DDCAR11T	I48979 Sequence 6
C	53	34	1.8	250	12	AB030158	X58193 D.discoiden
C	54	34	1.8	250	12	AB030158	AB030158 Pinus den

## ALIGNMENTS

RESULT	1		
LOCUS	AX077226	1912 bp	DNA
DEFINITION	Sequence	3	from Patent WO0105975.
ACCESSION	AX077226		
VERSION	AX077226.1	GI:13121820	
KEYWORDS			
SOURCE			
ORGANISM			
	Zea mays.		
	Zea mays		
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
	1 (bases 1 to 1912)		
	Mahajan,P.B.		
	Maize orthologues of bacterial ruvB:cdas and uses thereof		
	Patent: WO 0105975-A 3 25-JAN-2001;		
	PIONEER HI-BRED INTERNATIONAL, INC. (US)		
FEATURES			
source	location/Qualifiers		
	1..1912		
	/organism="Zea mays"		
	/db_xref="taxon:4577"		
	94..1461		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC32389.1"		
	/db_xref="GI:13121821"		
	/translation="MRIEEVQSTSKORIATHTHKGIDPANGMAITAIAGPFGQAA AREAGIADVMIRQKMKAGRAVLGAPGATGTALGIAOELGSKVPCPMQSEVY SSEVKKEIVEMERRAIGRIKENKEVEYGEVPELSPREESTGTGAKSISHVITS LKYVKGKOLKIDSTSYDALIREKVAAGVYITENSGAVNRVCDSFATYEDLEAE EYVPIPGEYVHKKEIIVQDVTTLHDDIDANADPQGGDDILSTMGQMMPRTEITTEKLR		

```

Db 901 GAAATCACCGGAAAACTACGCCAAGAAATTAATAGGTGGTAATATGATATATGATGAA 960
Qy 961 ggaattcagagcttgcaccgtgtgttttttcaatgaatggtccacatttggatc 1020
Db 961 ggaattcagagcttgcaccgtgtgttttttcaatgaatggtccacatttggatc 1020
Qy 1021 gaatttttcttaacttaacccgtcatcttgagagccattatcaccaatcgtatcct 1080
Db 1021 GAATGTTTCTTATCTTAACCGTGATGAGAGGCCATTATACCAATCGTATCTT 1080
Qy 1081 gctcaaaataggggaataatgtaataaggaagacatataatgaagtcacatgata 1140
Db 1081 GCTACAAATAGGGAATATGTAATGTAAGAGGAATGATATGACAAATGCAATGATGATA 1140
Qy 1141 ccggtgatacttctaagatagctgtgtatattatctggcagaagataatgacctactag 1200
Db 1141 CCGGTGATCTTCTAGATAGGCTGTGATTTATGACAGACATATGACCTACTAGAG 1200
Qy 1201 atgatacagatattgctatccagacaaagtggagagatgatatgataagaagt 1260
Db 1201 ATGATACAGATATTGCTATCCAGACAACTGAGAGAGATTGATATGATGAGAGAGT 1260
Qy 1261 ctgtctattatagcagagatcgacagacacatctttgagacatgctattcaattgata 1320
Db 1261 CTGTCTATTATTAGCGAGATCGACAGACAGACATCTTTGAGACATGCTATTCAATTGATA 1320
Qy 1321 tcacctccagctgtgtctcaaaactaatggagaagaagaatctgcaagctgatactc 1380
Db 1321 TCACCTCCAGCTGTGTCTCAAAAGCTAATGAGAGAGAAATCTGCAAGGCTGATCTC 1380
Qy 1381 gaggaaatcagtgagctctatcttgatgcaaatccctcagctcagctccaggaacaa 1440
Db 1381 GAGGAATCAGTGGGCTCTATTTGATGCCAAATCTCGGCTCGCTCCAGAGACAA 1440
Qy 1441 caagaaagataacatcacatgatttgcctcgtcgtggaagctcgaagaagatgta 1500
Db 1441 CAAGAAAGATACATCACCTGATTTGATCTCTGCTGGAAGTCTCGAAGAAATGTA 1500
Qy 1501 gttgcagctcgaaagtcatcctagtcattgatactgcttcaaggtcattagctactg 1560
Db 1501 GTTGCCAGCTCGAAAGTCACTAGTGCATGATCTGCTTCAAGGTTCAATAGTCTACTGG 1560
Qy 1561 tcttgagagcagacatttcggggggaacggtcgaatttggacgtccgtctgtgta 1620
Db 1561 TCTTGAGAGCGACATTTTGGGGGGAACGCTTGAATTTGCAAGTCCGTGCTGTGTTA 1620
Qy 1621 gttccaragaagactgtgtccgcacatctgctcgttcacgacgttccgtgattaga 1680
Db 1621 GTTCCARAGAAGACTGTGTCCGCGCATTTGCTCGTTACGCGCTGTGCGATTAGA 1680
Qy 1681 ttgtgcacggctgcaagaattgcgctgtgttttatacttgcacagctgcgcgaagt 1740
Db 1681 TTGTGCACCGGCTGCAAGAAATTTGCCGTGTGTTTATCTTGCATACGCTGCCGAGT 1740
Qy 1741 ctgctccacgggtgtattggtgcccgaacctatcttforaacatgagataagtagg 1800
Db 1741 CTGCTCCACGGGTGTATTGTTGGCCCGAACCTTATCTTGAACCATGGAATAGGATAGG 1800
Qy 1801 attctcaagaatgcaactgtcatgcttattatcttaactcaatgcatataaagcataag 1860
Db 1801 ATTCTTACAGAAATGCAACTGTGATGCTTATTATTTCTTAATGTCATTAACATTAAGC 1860
Qy 1861 aaatgttctcaaaemtwtataaaaaaataaaaaaataaaaaaataaaaaa 1912
Db 1861 AAATGTTCTTACAACMTWTAAAAAATAAAAAAATAAAAAAATAAAAAA 1912

```

## RESULT 2

```

AX077232 1869 bp DNA PAT 22-FEB-2001
LOCUS AX077232
DEFINITION Sequence 9 from Patent WO0105975.
ACCESSION AX077232

```

```

VERSION AX077232.1 GI:13121826
KEYWORDS
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1869)
REFERENCE
AUTHORS Mahajan, P. B.
TITLE Maize orthologues of bacterial ruvB:cdnas and uses thereof
JOURNAL Patent: WO 0105975-A 9 25-JAN-2001;
PIONEER Hi-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1. 1869
/organism="Zea mays"
/db_xref="taxon:4577"
64..1380
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC32392.1"
/db_xref="GI:13121827"
/translation="WRITEVSTSKKRIATHTTHIKIGLDANGMATAIAAGYQOK
MAGRAVLLAGPATGKTALAIQELGSKVPCFVWGSSEVKSKEVLEMFRR
IGLRIRKEKVEYESEVTELSPEAEESTGSAKISHYIISLKTVEKTKOLKDSIT
DALIKKEVAGDVYIIEANSAGVKRVKRGSEFAEYDEAEVYPIPEGVHKKETV
ODVTLHDIDANAOOGGODILSLMGOMKMRKEITEKLOELINKVYNNRIIDEGTAE
LYPVGLFIDEVHMDIECFSTYLNRLSEPSYITLATNRGICNVRGDTSPGITY
DLIDLVLIRIETGTETMIOILAIRAVEEIDIDESLAIUGETGQOTSURHAIQLI
SPASVSTNGRKEKICKADLEVSGLYIDAKSSARLLOBOBERIT"
BASE COUNT 539 a 373 c 486 g 471 t
ORIGIN
Query Match 56.5%; Score 1080; DB 10; Length 1869;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1130; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 337 ggcataccccagagctggcaggaagtcctcttctgctcatggttggatcagaagt 396
Db 256 GCGATAGCCCGAGAGCTCGGACGACAAAGTCCCTTCTGTCATGATGATGATGACAAAGG 315
Qy 397 tactctcgaggtcacaagaactgaggtgctgataagaaatttcgtagagctatagt 456
Db 316 TACTCTCGGAGGTCAAGAAACTGAGTGTGATGAGAAATTTCCGTAGACTATAGT 375
Qy 457 ttgcgtataaagaaacaagaaggttataagagagaggttactgaacttccacaga 516
Db 376 TTGCGTATTAAGGAACAAGAGGTTTATGAAGAGAGAGGTTACTGAACCTTCCACAGA 435
Qy 517 gaggctgaagatcacactgtgtgatatgataaagaacattagacatgataatacagctta 576
Db 436 GAGGCTGAGAGTACCACTGCTGGATATGCAAAAAGCATTAATGCAATGTCACGCTTA 495
Qy 577 aagactgttaagggactaagaactgaagttagatttctcaattatgatactctgtac 636
Db 496 AAGACTGTTAAAGGAGCTAAGCACTGAAGTATGCTTCAATTAATGATAGCTGTGATC 555
Qy 637 aaggaagaagtgagcagtgagggtatgatatcatcatcogaagaactagtggagcagtgaa 696
Db 556 AAGGAAGAAGCTGCGAGTGGGTATGATATACATTCGAAGCAAAATAGTGAGACACTGAAA 615
Qy 697 agagttgtagatgtgattcttcttcaagaatacgaactcgtgaagctgaagaagtatgt 756
Db 616 AGAGTTGGTAGATGTGATTCTTTTGTCTACAGAAATGAGATCTTGAAGCTGGAAGATGTGT 675
Qy 757 cctatccccaaggtgaagtcataaagaaagaatgtgtcagagatgtccactcat 816
Db 676 CCAATCCCAAGGTGAAGTCCATTAAGAAAAAAGAAATGTGTCAGAGATGTACACTTTCAT 735
Qy 817 gaccttgatgcagcaaatgtcaccacaaggtggcagaagatattgtccctatgggc 876
Db 736 GACCTTGATGCAGCAAAATGCTCAGCCACAGAGTGGCCAAAGATATTTGTCCTTATGGGC 795

```

QY 877 cagatgataaacacgaagaagctgaatccaccgaataaacctcgaagaataatgaag 936  
 |||||||  
 Db 796 CAGATGATGAACACGAAAGACTGAATCACCGAAATCAACGCAAAATTAATAG 855  
 QY 937 gtgttaaatgatatcgaatgaagaatgacgaagctgtacacctgtgttcatt 996  
 |||||||  
 Db 856 GTGTAAATGATATTCGATGAAGCAATTGCGAGCTGTACCTGGTGTTCATT 915  
 QY 997 gatgaagtcacatgtgtgatalcgaatgtttcttacttaacccgtgcatgtgagag 1056  
 |||||||  
 Db 916 GATGAGGTCCACATGTTGGATATCGAATGTTTCTTATCTTAACCGCTCATTTGGAGACC 975  
 QY 1057 ccaattacacaaatcgtgaactgtctacaataagggaatgtgaatgaagaact 1116  
 |||||||  
 Db 976 CCAATTATCCCAATCGTGAATCTGCTACAAATAGGGGAATGATATTAAGAGAACT 1035  
 QY 1117 gataatgaagtcacatggtataccggtgtgactcttagatagctgtgttatttcg 1176  
 |||||||  
 Db 1036 GATATGACAAAGTCCACATGATATACCGGTGATCTCTGATATGCTGTATTTATTCG 1095  
 QY 1177 acagaacacatagtgacctactgagatgatacagatatgtgctacccgacaaatgag 1236  
 |||||||  
 Db 1096 ACAGAGACATATGCGCTTACGTAGATGATACAGATATTTGCTATCCGACACAAAGTGAG 1155  
 QY 1237 ggaattgatatgatagaagaagctgtctactttagagcagaatgagacagacatct 1296  
 |||||||  
 Db 1156 GAATTTGATATGATGATGAAGAAAGTCTTGCTTATTTAGGGGAGATCGGACAGACATCT 1215  
 QY 1297 ttgagacatgctatcaaatgtatatacactgacagcgtgtgtctcaagaactatgaga 1356  
 |||||||  
 Db 1216 TTGAGACATGCTATTCATATGATATCACCGCCGCGTGTCAAAAGACTATGAGAA 1275  
 QY 1357 gaaaaaatcgcgaagctgtgctcgaagaagtcagtgagctcatttgatgccaaatcc 1416  
 |||||||  
 Db 1276 GAGAAATCTGCAAGGCTGATCTCGAGAGATTAGTGGCTTATTTGGATGCCAATCC 1335  
 QY 1417 tcggtctgctgtctcagagagcaagaagaatacatcaactagattgg 1467  
 |||||||  
 Db 1336 TCGGCTGCGCTGCCAGGAGCAAAAGATACATCATCAGATTGG 1386

RESULT 3  
 AX077224 1845 bp DNA PAT 22-FEB-2001  
 LOCUS AX077224 Sequence 1 from Patent WO0105975.  
 DEFINITION AX077224  
 ACCESSION AX077224  
 VERSION AX077224.1 GI:13121818  
 KEYWORDS  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 1845)  
 AUTHORS Mahajan, P. B.  
 TITLE Maize orthologues of bacterial ruvb:cdnas and uses thereof  
 JOURNAL Patent: WO 0105975-A 1 25-JAN-2001;  
 PIONEER HI-BRED INTERNATIONAL, INC. (US)

## FEATURES

source  
 1. 1845  
 /organism="Zea mays"  
 /db\_xref="taxon:4577"  
 85. 1452  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC32388.1"  
 /db\_xref="GI:13121819"  
 /translation="MRIEVOSTSKRORITATHTHIKGLDNGMAIALAAGFVGQSA  
 AREAGIADMITROKMGRAVLGVPAATGCTALALIAQELGSKVPCPMVGSEVY  
 SGEVKTETVIMNFRATIGRLIKENKVEGEGTALSPBEASGTGAKSISHTIS  
 LKTVGCTROKIDSSIDVALIKERKAVANDVITYEANSQAVRVGCDSPATYELEA  
 EYVPIKGEVHKKEIVDVTIHDIDANADPOGGODILSLMGOMKRRKREITTEKIL  
 QEINRVNRY IDEGIAELVGVLFIDEVHMDIECFSTLNRLAESPLSIVLTNRG

ICNVKGTDMTSPHGIPIVDLRLVITRTETGYCTEMIOLIAIQAQVEDIDIMDESLEY  
 LGEIGQOTSLRHAHQILISIPASVSKINRGRIKIDKALEBVSGLYLAKKSARLLQEQO  
 ERYIT

Query Match 30.1%; Score 576; DB 10; Length 1845;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1426; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

BASE COUNT 530 a 378 c 490 g 445 t 2 others  
 ORIGIN  
 QY 84 gtgcgagcgaatgaagatcagagagtgacgtcagactcgaagaagcagcgcac 143  
 |||||||  
 Db 75 GTTGGCGGCGATGAGGATCGAGAGGTGCTCACTCGAAGAAAGCGCATCGCAC 134  
 QY 144 ccaacccacatcaaggagactcggctcgcagcgaatgagatgagatgagtgagcgc 203  
 |||||||  
 Db 135 CCACACCCACATCAAGGAGCTCGCTCGACGCCAATGGATGGCGATTGGTGGCGGC 194  
 QY 204 ggggttcgtggccagagcgcgcgcgagcgagcgcgagcgtgagcgtgagcgtgagc 263  
 |||||||  
 Db 195 GGGGTTGCTGGGCGCAGTGGCGCGCGCGCGAGCGCGCGCTGGCGGTCCACATGATTGG 254  
 QY 264 ccaagaagaatgagc 323  
 |||||||  
 Db 225 CCAGAAAAAGATGGCGCGCGCGCGCGCTCTGCTGCTGGCTCGCGCGCGCGCAAGAC 314  
 QY 324 ggcgcctagcgtcgc 383  
 |||||||  
 Db 315 GCGCGTACCGCTCGGCAATAGCCAGGAGCTCGGAGCAAGGCTCCCTTCTGCTTATGGT 374  
 QY 384 aggatcaagaatgtagctcctcgcagagtgcaagaagaactgagtgctgatagaaatlcgc 443  
 |||||||  
 Db 375 AGGATCAAGATGATCTCTCGAGGTCAAGAAACAGAGTGTCTGATGAAATTTCCG 434  
 QY 444 tagagctaaagtttggtatataaagaagaagaagttatagaagagagtgactga 503  
 |||||||  
 Db 435 TAGACCTATGAGTGTGCTTAAAGGAAACAAAGAGGTTTATTAAGAGGTTACTGA 494  
 QY 504 acttccccagaagaagctgtgagatgatacactggtgtgatatgataaagaactgaacatgc 563  
 |||||||  
 Db 495 ACTTTCCTCCAGAGAGGCTGAGAGTACACTGCTGATATGCAAAAGCATTTAGCCATGT 554  
 QY 564 aatcacagcttaaaagactgtttaaaggagactaaagcaactgaagttagatcttcaattta 623  
 |||||||  
 Db 555 AATCATGAGCTTAAAGACTGTTAAAGGAGCTAAGCACTAAGTTAGATTCTCAATTGA 614  
 QY 624 tgaatccttgatcaagaagaagtgagcagtggtgtgagtgtatatacatcgaaagcgaatag 683  
 |||||||  
 Db 615 TGATGCTCTGATCAAGGAAAGGTGGCGAGTGGGTGATGTTATTAACATTGAAGCAAAATAG 674  
 QY 684 tggagcagtgaaaagagttgtgtagatgtagtctcttgcctacagaatacgaatcttgaagc 743  
 |||||||  
 Db 675 TGGACAGTGAAGAAAGTGGTGTAGTGTGATTTCTTTTGCTACAAATACGATCTTGAAGC 734  
 QY 744 tgaagaatgtagtcttccatccccaaaagtgaaagttccataaagaagaagaatgtgcaaga 803  
 |||||||  
 Db 735 TGAAGAAATATGTTCTTATCCCAAGGAGTGAAGTCCATTAAGAAAAAATAGTGCAGGA 794  
 QY 804 tgtcaacttcaatgaaccttgatgcaagaatgagcagcagcgaagaagtggtgccaagtatttc 863  
 |||||||  
 Db 795 TGTCACTTCAATGACCTTGTATGACGCAAAATGCTACGCCAAGAGTGGCCCAAGATATTTT 854  
 QY 864 gtccctttagggccagatgtagtgaacccagaaagactgaaatcccgaaaaacacagcga 923  
 |||||||  
 Db 855 GTCCCTTTATGGGCCAGATGATGAACCAAGCAAGATGATATCCGAAAAAATAGTACGCCA 914  
 QY 924 agaaatataaagtgtagaataatgatalatacgaatgaagaatgtgcagagcttgtaacctgg 983  
 |||||||  
 Db 915 AGAAATTAATTAAGGTGTAATATATATATGATGAAGGAATTCAGAGCTTGTACTCGG 974  
 QY 984 tgtttgttcaatgtgaggtcccaatgttgatatacgaatgtttcttatacctaacgcg 1043  
 |||||||

```

Db 975 TGTATTATTCATGACGAGGTCCACATGTTGATATGCAATGTTTCTTATCTTAACCG 1034
Qy 1044 tgcattggagagcccatatcatccaatcgtgatacttgatacaaatagggaatatgtaa 1103
Db 1035 TGCATTGGAGAGCCCATTAACCAATTTGTGATCTTGCTACGCAATATGGGAATATGCTAA 1094
Qy 1104 tctaaaggaactatgatgcaaatgacacatgfatccagtgatgcttcttaataagct 1163
Db 1095 TGTAGAGGAACGTATATGACAAAGTCCACATGATATACCAAGTATCTTATAGATTAGGTT 1154
Qy 1164 ggtattatttcgagacagacataatgagccctactgagatgatacagatattgctatccg 1223
Db 1155 GGTATTATTTCGACACAGACATATGSCCTACTGACATATATACAGATATTGGCTATACCG 1214
Qy 1224 agcaaatgagagagatgataatgataatgataatgataatgataatgataatgataatg 1283
Db 1215 AGCAACAGTGGAGGACATTTATATGATGTAAGAAAGCTTGTATTTAGGCGAGATCGG 1274
Qy 1284 acaagcaacatcttggagacatgctatgataatgataatgataatgataatgataatgataatg 1343
Db 1275 ACACGACACATCTTTAAGACATGCTATTCAATTGATATACCTGCGCCAGCTGCTCATAA 1334
Qy 1344 gactaatgagagagagagagagagagagagagagagagagagagagagagagagagagagag 1403
Db 1335 GACTAATGGAAGAGAGAAATATGCAAGGCTGATCTCGAGGAAGTCAAGTGGCTATATT 1394
Qy 1404 ggaatgcaaatctcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1463
Db 1395 GGATGCCAAATCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 1454
Qy 1464 ttgagatctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1523
Db 1455 TTGGATTCACCTGCTGCTGGAAGTCTCGAAGAGATGTAGTTGCCACCTCGAAAGTCACTTA 1514
Qy 1524 gtcg 1526
Db 1515 GTG 1517

RESULT 4
LOCUS AX077228 1886 bp DNA PAT 22-FEB-2001
DEFINITION Sequence 5 from Patent WO0105975.
ACCESSION AX077228
VERSION AX077228.1 GI:13121822
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
TITLE Panicoidae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 1886)
PIONEER Maize orthologues of bacterial ruvb:cdnas and uses thereof
Patent: WO 0105975-A 5 25-JAN-2001;
LOCATION/Qualifiers
1.1886
/organism="Zea mays"
/db_xref="taxon:4577"
82..1449
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC32390.1"
/db_xref="GI:13121823"
/translation="MRIEVDSTSKKRIATHTIHKGLDANGMSMPLAAGVGOAA
AREAGLAVDMIRQKMAGRALLAGPAPATGALALGIAOELGSKVPCPMVGSSEV
SSEVKTEVLMENFRRAIGLRIKENKEVEYELSPERASTGGVAKSISHVILG
LKVTKTKOLKIDPSIYDALIKERKAVGVITYEANSQAVKRVGRCDSPATEYDLAE
EYVPIPKGEVHKKEIYODVTLHDIDANAOPGODIISLMGOMKPKRTEITELR
QETNKVYNRIDEGIALVGVFLIDEVHMDIECSYLNRALSPISPLVILATNR
ICNVKGTDMTSPHGIPTVLDRLVITRTTYGPTMTQIATRAOYEIDIDESLAY
LGEIGQOTSIRHAIQLLSPASVYAKTNGREKMKADLEVSGLYLDAKSSARLLQEOQ

```

```

BASE COUNT 524 a 408 c 518 g 436 t
ORIGIN
Query Match 4.9%; Score 94; DB 10; Length 1886;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 406 gaggtcaagaacacgaggtgctgataatgataatgataatgataatgataatgataatgataatg 465
Db 481 GAGGTCAAGAAACCTGAGGTGCTGATGAAATTTCCGATAGCTATAGGTTGGCTATA 540
Qy 466 aaggaacacacagaggttataagagagaggtta 499
Db 541 AAGGAAACAAAGAGGTTATGAGAGAGAGGTTA 574

RESULT 6
LOCUS AP003271 171809 bp DNA HTG 21-FEB-2001
DEFINITION
ACCESSION AP003271
VERSION AP003271.1 GI:13121822
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
TITLE Panicoidae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 1898)
PIONEER Maize orthologues of bacterial ruvb:cdnas and uses thereof
Patent: WO 0105975-A 7 25-JAN-2001;
LOCATION/Qualifiers
1.1898
/organism="Zea mays"
/db_xref="taxon:4577"
166..1536
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC32391.1"
/db_xref="GI:13121825"
/translation="MRIEVDSTSKKRIATHTIHKGLDANGMSMPLAAGVGOAA
AREAGLAVDMIRQKMAGRALLAGPAPATGALALGIAOELGSKVPCPMVGSSEV
SSEVKTEVLMENFRRAIGLRIKENKEVEYELSPERASTGGVAKSISHVILG
LKVTKTKOLKIDPSIYDALIKERKAVGVITYEANSQAVKRVGRCDSPATEYDLAE
EYVPIPKGEVHKKEIYODVTLHDIDANAOPGODIISLMGOMKPKRTEITELR
QETNKVYNRIDEGIALVGVFLIDEVHMDIECSYLNRALSPISPLVILATNR
ICNVKGTDMTSPHGIPTVLDRLVITRTTYGPTMTQIATRAOYEIDIDESLAY
LGEIGQOTSIRHAIQLLSPASVYAKTNGREKMKADLEVSGLYLDAKSSARLLQEOQ

```

DEFINITION Oryza sativa chromosome 1 clone P0506B12, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, in ordered pieces.  
 ACCESSION AP003271  
 VERSION AP003271.1 GI:13027301  
 KEYWORDS HTG; HTGS\_PHASE2.  
 SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:P0506B12.  
 ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae;  
 Oryza;  
 1 (sites)  
 Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 Oryza sativa nipponbare(GA5) genomic DNA, chromosome 1, PAC  
 clone:P0506B12  
 JOURNAL Published Only in Database (2001) In press  
 REFERENCE 2 (bases 1 to 171809)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of  
 Agricultural Resources, Rice Genome Research Program; Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces is believed  
 to be correct as given, however the sizes of the gaps between them  
 are based on estimates that have provided by the submitter. This  
 sequence will be replaced by the finished sequence as soon as it is  
 available and the accession number will be preserved.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 FEATURES  
 source  
 1..171809  
 /organism="Oryza sativa"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /chromosome="1"  
 /clone="P0506B12"  
 BASE COUNT 49272 a 36863 c 37265 g 48309 t 100 others  
 ORIGIN  
 Query Match 2.1%; Score 40; DB 83; Length 171809;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 688 gcagtgaaagaggttgtagatgattcttctacagaatacagatctgaagctga 746  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 7696 CGAGTGAAGAGGTTGTGATGATCTTTTGTGCTACAGATACGATCTGACGCTGA 7638  
 RESULT 7  
 LOCUS CNS06LVV 1173 bp DNA STS 10-JAN-2001  
 DEFINITION T7 end of clone XAT0A001H09 of library XAT0AA from strain CBS 4311  
 of Saccharomyces servazzii, sequence tagged site.  
 ACCESSION AL404705  
 VERSION AL404705.1 GI:12166507  
 KEYWORDS STS.  
 SOURCE Saccharomyces servazzii.  
 ORGANISM Saccharomyces servazzii.  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 1 (bases 1 to 1173)  
 Casaregola,S., Lepingle,A., Bon,E., Neuvéglise,C., Nguyen,H.,  
 Artiguenave,F., Wincker,P. and Galliard,J.C.  
 TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 7.  
 JOURNAL Saccharomyces servazzii  
 PUBMED FEBS Lett. 487 (1), 47-51 (2000)  
 1152882  
 REFERENCE 2 (bases 1 to 1173)

AUTHORS Souclet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
 Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
 Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S.,  
 Saurin,M., Tekala,F., Toffano-Nioche,C., Weslowski-Louvel,M.,  
 Wincker,P. and Weissensbach,J.  
 TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of  
 yeast species for molecular evolution studies(1)  
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
 PUBMED 1152876  
 REFERENCE 3 (bases 1 to 1173)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,  
 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT This STS is part of a random genomic sequencing program of thirteen  
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
 5 kb were prepared and both extremities were sequenced. See  
 keywords for description of this sequence and for the sequence of  
 the other extremity of this insert.  
 FEATURES  
 source  
 1..1173  
 /organism="Saccharomyces servazzii"  
 /strain="CBS 4311"  
 /db\_xref="taxon:27293"  
 /clone="XAT0A001H09"  
 /clone\_1bp="XAT0AA"  
 /note="end : T7"  
 STS 1..1173  
 /note="part of mitochondrial DNA"  
 /evidence=not\_experimental  
 BASE COUNT 405 a 173 c 64 g 398 t 133 others  
 ORIGIN  
 Query Match 1.8%; Score 35; DB 53; Length 1173;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-09;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1878 wtaaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 1912  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1104 WTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1138  
 RESULT 8  
 LOCUS I29928 55 bp DNA PAT 06-FEB-1997  
 DEFINITION Sequence 41 from patent US 5578468.  
 ACCESSION I29928  
 VERSION I29928.1 GI:1820719  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 1 (bases 1 to 55)  
 AUTHORS Pickup,D.J., Patel,D. and Antczak,J.B.  
 TITLE Site-specific RNA cleavage  
 JOURNAL Patent: US 5578468-A 41 26-NOV-1996;  
 FEATURES Location/Qualifiers  
 source  
 1..55  
 /organism="unknown"  
 BASE COUNT 47 a 2 c 1 g 5 t  
 ORIGIN  
 Query Match 1.8%; Score 34; DB 10; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 ||||||||||||||||||||||||||||  
 Db 15 TAAAAAAAAAAAAAAAAAAAAAAAAA 48

RESULT 9

LOCUS S76508 71 bp DNA PRI 07-MAY-1993

DEFINITION Dis8 (B) [hypervariable minisatellite MS32] [human, individual

AS89, Genomic, 71 nt].

ACCESSION S76508

VERSION S76508.1 GI:243383

KEYWORDS human individual AS89.

SOURCE

human individual AS89.

ORGANISM

Homo sapiens

REFERENCE

Monckton, D.G. and Jeffreys, A.J.

AUTHORS

Minisatellite 'isocollide' discrimination in pseudohomozygotes by

TITLE

single molecule PCR and variant repeat mapping

JOURNAL

Genomics 11 (2), 465-467 (1991)

MEDLINE

92120671

REMARK

GenBank staff at the National Library of Medicine created this

entry [NCBI g1bbsq 76508] from the original journal article.

This sequence comes from Figure 2.

Location/Qualifiers

1..71

source

/db\_xref="taxon:9606"

gene

/gene="Dis8"

BASE COUNT

66 a 0 c 0 g 5 t

ORIGIN

Query Match 1.8%; Score 34; DB 97; Length 71;

Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taaaaaaaaaaaaaaaaaaaaaaaaa 1912

||||||||||||||||||||||||||

Db 2 TAAAAAAAAAAAAAAAAAAAAAAAAA 35

RESULT 10

LOCUS S76510 72 bp DNA PRI 07-MAY-1993

DEFINITION Dis8 (B) [hypervariable minisatellite MS32] [human, individual

AS89, Genomic, 72 nt].

ACCESSION S76510

VERSION S76510.1 GI:243381

KEYWORDS human individual MACH.

SOURCE

human individual MACH.

ORGANISM

Homo sapiens

REFERENCE

Monckton, D.G. and Jeffreys, A.J.

AUTHORS

Minisatellite 'isocollide' discrimination in pseudohomozygotes by

TITLE

single molecule PCR and variant repeat mapping

JOURNAL

Genomics 11 (2), 465-467 (1991)

MEDLINE

92120671

REMARK

GenBank staff at the National Library of Medicine created this

entry [NCBI g1bbsq 76510] from the original journal article.

This sequence comes from Figure 2.

Location/Qualifiers

1..72

source

/db\_xref="taxon:9606"

gene

1..72

/gene="Dis8"

/allele="B"

BASE COUNT

62 a 0 c 0 g 10 t

ORIGIN

Query Match

1.8%; Score 34; DB 97; Length 72;

Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taaaaaaaaaaaaaaaaaaaaaaaaa 1912

||||||||||||||||||||||||||

Db 1 TAAAAAAAAAAAAAAAAAAAAAAAAA 34

RESULT 11

LOCUS A08915 73 bp DNA PAT 02-SEP-1993

DEFINITION H.sapiens (haplotype 3, allele MS32, isolate French, serial number

8) minisatellite sequence.

ACCESSION A08915

VERSION A08915.1 GI:411837

KEYWORDS human.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Monckton, D.G. and Jeffreys, A.J.

AUTHORS

Extended nucleotide sequences

TITLE

Patent: EP 0370719-A 96 30-MAY-1990;

JOURNAL

IMPERIAL CHEMICAL INDUSTRIES PLC

location/Qualifiers

1..73

source

/db\_xref="taxon:9606"

BASE COUNT

69 a 0 c 0 g 4 t

ORIGIN

Query Match

1.8%; Score 34; DB 9; Length 73;

Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taaaaaaaaaaaaaaaaaaaaaaaaa 1912

||||||||||||||||||||||||||

Db 1 TAAAAAAAAAAAAAAAAAAAAAAAAA 34

RESULT 12

LOCUS A08908 77 bp DNA PAT 02-SEP-1993

DEFINITION H.sapiens (haplotype 3, allele MS32, isolate Mormon, serial number

10) minisatellite sequence.

ACCESSION A08908

VERSION A08908.1 GI:411830

KEYWORDS human.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Monckton, D.G. and Jeffreys, A.J.

AUTHORS

Extended nucleotide sequences

TITLE

Patent: EP 0370719-A 91 30-MAY-1990;

JOURNAL

IMPERIAL CHEMICAL INDUSTRIES PLC

location/Qualifiers

1..77

source

/db\_xref="taxon:9606"

BASE COUNT

72 a 0 c 0 g 5 t

ORIGIN

Query Match 1.8%; Score 34; DB 9; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 |||  
 Db 5 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 38

RESULT 13  
 AX099452 79 bp DNA PAT 02-APR-2001  
 LOCUS Sequence 92 from Patent WO0119988.  
 ACCESSION AX099452  
 VERSION AX099452.1 GI:13538555  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 79)  
 AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Collins-Racie,L.A., Evans,C.,  
 Merberg,D., Treacy,M., Bowman,M.R., Spaulding,V. and Agostino,M.J.  
 TITLE Secreted proteins and polynucleotides encoding them  
 JOURNAL Patent: WO 0119988-A 92 22-MAR-2001;  
 FEATURES Location/Qualifiers  
 source 1..79  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 64 a 2 c 3 g 10 t  
 ORIGIN

Query Match 1.8%; Score 34; DB 10; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 |||  
 Db 27 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

RESULT 14  
 A08907 80 bp DNA PAT 02-SEP-1993  
 LOCUS H.sapiens (haplotype 3, allele MS32, isolate English, serial number  
 DEFINITION 11) minisatellite sequence.  
 ACCESSION A08907  
 VERSION A08907.1 GI:411829  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 80)  
 AUTHORS Jeffreys,A.J.  
 TITLE Extended nucleotide sequences  
 JOURNAL Patent: EP 0370719-A 90 30-MAY-1990;  
 IMPERIAL CHEMICAL INDUSTRIES PLC  
 FEATURES Location/Qualifiers  
 source 1..80  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 72 a 0 c 0 g 8 t  
 ORIGIN

Query Match 1.8%; Score 34; DB 9; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 |||  
 Db 7 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 40

RESULT 15  
 AX098165/c 83 bp DNA PAT 30-MAR-2001  
 LOCUS Sequence 77 from Patent WO0118542.  
 ACCESSION AX098165  
 VERSION AX098165.1 GI:13515205  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 83)  
 AUTHORS Lee,J., Thompson,P. and Lillie,J.  
 TITLE Identification, assessment, prevention, and therapy of ovarian  
 cancer  
 JOURNAL Patent: WO 0118542-A 77 15-MAR-2001;  
 Millennium Predictive Medicine, Inc. (US)  
 FEATURES Location/Qualifiers  
 source 1..83  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="All occurrences of n indicate any nucleotide"

BASE COUNT 12 a 10 c 13 g 47 t 1 others  
 ORIGIN

Query Match 1.8%; Score 34; DB 10; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 |||  
 Db 64 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 31

Search completed: November 13, 2001, 09:53:05  
 Job time: 6926 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 08:44:44 ; Search time 144.47 Seconds  
(without alignments)  
8310.015 Million cell updates/sec

Title: US-09-589-510-3

Perfect score: 1912  
Sequence: 1 acccaagcttcgcaattt.....aaaaaaaaaaaaaaaaaa 1912

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq\_0601.\*  
1: /SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1912	100.0	1912	22	MAIZE RuVB ortholog
2	1080	56.5	1865	22	MAIZE RuVB ortholog
3	576	30.1	1845	22	MAIZE RuVB ortholog
4	94	4.9	1886	22	MAIZE RuVB ortholog
5	94	4.9	1898	22	MAIZE RuVB ortholog
6	34	1.8	69	18	3' portion of CDNA
7	34	1.8	69	18	Human secreted pro
8	34	1.8	79	19	Human novel clone
9	34	1.8	79	19	Human novel secret
10	34	1.8	80	19	Oligonucleotide se
11	34	1.8	81	21	Human secreted pro

12	34	1.8	85	19	AAV05720	Nucleotide sequenc
13	34	1.8	86	12	AAO11760	Self-complementary
14	34	1.8	89	21	AACT8380	Human cancer assoc
15	34	1.8	90	12	AAO11762	Self-complementary
16	34	1.8	90	21	AAFL1895	Lung cancer associ
17	34	1.8	95	21	AAFL16473	Human prostate can
18	34	1.8	98	20	AAAX00174	Porcine reproducit
19	34	1.8	99	18	AAAT91300	Human M97-2 secret
20	34	1.8	101	18	AAV00420	3' fragment of clo
21	34	1.8	111	21	AAAC98625	Human colon cancer
22	34	1.8	114	17	AAAT39472	Growth regulatory
23	34	1.8	114	19	AAV37490	Human growth regul
24	34	1.8	120	14	AAO52429	Human subtelomeric
25	34	1.8	120	16	AAO87011	Subtelomeric CDNA
26	34	1.8	120	19	AAAT96338	Subtelomeric CDNA
27	34	1.8	140	18	AAAT76781	Staphylococcus aur
28	34	1.8	141	16	AAAT06012	Immunodominant fra
29	34	1.8	144	19	AAV21239	Homo sapiens clone
30	34	1.8	144	20	AAAT77331	Human secreted pro
31	34	1.8	191	17	AAAT13475	Capture probe for
32	34	1.8	198	21	AAAG4592	Nucleotide sequenc
33	34	1.8	201	18	AAV00418	3' fragment of clo
34	34	1.8	202	21	AAAC98563	Human colon cancer
35	34	1.8	208	20	AAAT14464	Human gene express
36	34	1.8	208	20	AAAT7547	Murine secreted pr
37	34	1.8	216	19	AAV09114	3' nucleotide sequ
38	34	1.8	216	20	AAAT77545	Murine secreted pr
39	34	1.8	216	21	AAAT59487	Sequence determin
40	34	1.8	222	14	AAO46071	Downstream sequenc
41	34	1.8	222	22	AAAC86777	Human lung tumour
42	34	1.8	229	22	AAAC86277	Human colon cancer
43	34	1.8	231	21	AAAC98505	Human colon cancer
44	34	1.8	234	21	AAAC98428	Human prostate can
45	34	1.8	236	21	AAFL6372	

#### ALIGNMENTS

RESULT 1	
AAO02567	
ID	AAO02567 standard; cDNA; 1912 bp.
XX	
AC	AAO02567;
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Maize RuVB orthologue #2 CDNA.
XX	
XX	Maize; RuVB orthologue; branch migration; heteroduplex extension;
KW	homologous recombination; transformation; transgenic plant; ss.
KM	
XX	
OS	Zea mays.
XX	
PH	Key
FT	94..1461
FT	/*tag= a
FT	/product= "Maize RuVB orthologue protein #2"
XX	
PN	W0200105975-A1.
XX	
PD	25-JAN-2001.
XX	
PF	13-JUN-2000; 2000WO-US16271.
XX	
PR	16-JUL-1999; 99US-0144112.
XX	
PA	(PTON-) PIONEER HI-BRED INT INC.
XX	
PI	Mahajan PB;
XX	
XX	WPI: 2001-159537/16.
DR	P-PSDB; AAY72563.



Db 1681 ttggtaccggtgacgaattccgtgtgtttttattctctcatcggtccggagt 1740  
 QY 1741 ctgctccacgggttgatgtgcccgaacctattctgttaacctgataatgataag 1800  
 Db 1741 ctgctccacgggttgatgtgcccgaacctattctgttaacctgataatgataag 1800  
 QY 1801 attcttaacgaatgcaactgtgcttatttatttcttaaatgltccataaagcataag 1860  
 Db 1801 attcttaacgaatgcaactgtgcttatttatttcttaaatgltccataaagcataag 1860  
 QY 1861 aatgtttctacaacmtwtaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 Db 1861 aatgtttctacaacmtwtaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 RESULT 2  
 AAD02570  
 ID AAD02570 standard; cDNA; 1869 BP.  
 AC AAD02570;  
 DT 02-MAY-2001 (first entry)  
 DE Maize RuVB orthologue #5 cDNA.  
 KW Maize; RuVB orthologue; branch migration; heteroduplex extension;  
 KM homologous recombination; transformation; transgenic plant; ss.  
 OS Zea mays.  
 FT Key Location/Qualifiers  
 FT CDS 64..1380  
 FT /tag= a  
 FT /product= "Maize RuVB orthologue protein #5"  
 PN W0200105975-A1.  
 PD 25-JAN-2001.  
 PE 13-JUN-2000; 2000MC-US16271.  
 PR 16-JUL-1999; 99US-0144112.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 MA Mahajan PB;  
 WP1: 2001-159537/16.  
 DR P-PSDB: AAY72566.  
 PT Novel maize RuVB nucleic acid useful for modulating levels of maize  
 PT RuVB in plants, as probes or amplification primers in the detection,  
 PT quantitation or isolation of gene transcripts -  
 PS Claim 1; Page 80-82; 87pp; English.  
 XX The present sequence is a Zea mays RuVB orthologue #5 cDNA. RuVB along  
 CC with RuVB catalyses the branch migration process, also known as  
 CC heteroduplex extension, in homologous recombination. RuVB is used for  
 CC the control of homologous recombination or transformation efficiency in  
 CC transgenic plants. The RuVB nucleotide may be used as probes or  
 CC amplification primers for detecting, quantifying or isolating gene  
 CC transcripts, in detecting deficiencies in the mRNA level during screening  
 CC for desired transgenic plants, for detecting gene mutations, for  
 CC monitoring upregulation of expression or changes in enzyme activity, for  
 CC detecting any number of allelic variants, orthologues or paralogues of  
 CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
 CC also be used for recombinant expression of its encoded polypeptide, or  
 CC for use as immunogen in preparing and/or screening of antibodies, and in  
 CC sense or antisense suppression of one or more genes in a host cell,  
 CC tissue or plant. The RuVB proteins may be used in assays to agonise or  
 CC antagonise the enzyme function, or as immunogens or antigens for  
 CC screening antibodies.

XX  
 SQ Sequence 1869 BP; 539 A; 373 C; 486 G; 471 T; 0 other;  
 Query Match 56.5%; Score 1080; DB 22; Length 1869;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1130; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 337 ggcataagcccgagggatctgcgcagcaagaagttcccttctctctatgtagagtagaagatg 396  
 Db 256 ggcataagcccgagggatctgcgcagcaagaagttcccttctctctatgtagagtagaagatg 315  
 QY 397 tactctcgaggggtcaagaactgaagtgctgtagtgaagaaatlccgttagagctatagtt 456  
 Db 316 tactctcgaggggtcaagaactgaagtgctgtagtgaagaaatlccgttagagctatagtt 375  
 QY 457 ttgcgtataaaggaatacaagaaggttataagaaagagaggttactgaacttccccagaa 516  
 Db 376 ttgcgtataaaggaatacaagaaggttataagaaagaggttactgaacttccccagaa 435  
 QY 517 gaggtctgaggtacgaactggttgatatagcaaaagcatagccatgatacatcagctta 576  
 Db 436 gaggtctgaggtacgaactggttgatatagcaaaagcatagccatgatacatcagctta 495  
 QY 577 aagactgttaaaaggagactgaagcaactgaagttagatttcttaattatgagtcttgatc 636  
 Db 496 aagactgttaaaaggagactgaagcaactgaagttagatttcttaattatgagtcttgatc 555  
 QY 637 aaggaagaagtggtcagtggtgtagtattatatacgcgaagcaaatagtgtgagcagtgaaa 696  
 Db 556 aaggaagaagtggtcagtggtgtagtattatatacgcgaagcaaatagtgtgagcagtgaaa 615  
 QY 697 aaggttgtagtagttagattcttcttctcaagaatacagctcttgaagcgtgaagatgatt 756  
 Db 616 aaggttgtagtagttagattcttcttctcaagaatacagctcttgaagcgtgaagatgatt 675  
 QY 757 cctatcccccaagggtgaagtcctataagaataaagaatgtgtcagagatgtcaacttcatt 816  
 Db 676 cctatcccccaagggtgaagtcctataagaataaagaatgtgtcagagatgtcaacttcatt 735  
 QY 817 gaacttgatgacgaataatgtctcagcccaaggttgccaaagattttgtccctatagggc 876  
 Db 736 gaacttgatgacgaataatgtctcagcccaaggttgccaaagattttgtccctatagggc 795  
 QY 877 cagatgatagaacccacgaagaactgaatcacccgaagaactacgcgaagaataataaag 936  
 Db 796 cagatgatagaacccacgaagaactgaatcacccgaagaactacgcgaagaataataaag 855  
 QY 937 gtggtataatagatatatcagaagaagatgtcagaagctgtgaactgtgtgtttgttcatt 996  
 Db 856 gtggtataatagatatatcagaagaagatgtcagaagctgtgaactgtgtgtttgttcatt 915  
 QY 997 gatgaagttccacatgttgatatacgaatgttttcttcttaaccggtcatgtgaagtc 1056  
 Db 916 gatgaagttccacatgttgatatacgaatgttttcttcttaaccggtcatgtgaagtc 975  
 QY 1057 ccaattacccaatcgtgtgtactgtgtacaataagggagataatgttaattgaaggaagact 1116  
 Db 976 ccaattacccaatcgtgtgtactgtgtacaataagggagataatgttaattgaaggaagact 1035  
 QY 1117 gatatacgaatcacaatgtagtacccgtgtgactctctagataagctgtgtatttcgg 1176  
 Db 1036 gatatacgaatcacaatgtagtacccgtgtgactctctagataagctgtgtatttcgg 1095  
 QY 1177 acgagaacatattggccctatgagatgaagaagattgtgtctatcccgagcaagaattgag 1236  
 Db 1096 acgagaacatattggccctatgagatgaagaagattgtgtctatcccgagcaagaattgag 1155  
 QY 1237 gaattgatatatgaataaagaagcttcttattttagggagatcggagaagacagatct 1296  
 Db 1156 gaattgatatatgaataaagaagcttcttattttagggagatcggagaagacagatct 1215  
 QY 1297 ttgagacatgtctatcaattgatatcaactgtccagcgtgtgtctcaaaagcataatgaaaga 1356

```

Db      1216 ttgagatgcttcttcaatgatataccgacgctggtcccaagaactatggaaga 1275
QY      1357 gagaaatctgcaaggctgactcgcgagaagtcagtggtgctctatttgatgccaatcc 1416
Db      1276 gagaaatctgcaaggctgactcgcgagaagtcagtggtgctctatttgatgccaatcc 1335
QY      1417 tcggtcgcgtgtctccagagacacaagaagatacatcattggtg 1467
Db      1336 tcggtcgcgtgtctccagagacacaagaagatacatcattggtg 1386

RESULT  3
AAD02566
ID      AAD02566 standard; cDNA; 1845 BP.
XX
AC      AAD02566;
XX
DT      02-MAY-2001 (first entry)
XX
DE      Maize Ruvb orthologue #1 cDNA.
XX
KW      Maize; Ruvb orthologue; branch migration; heteroduplex extension;
OS      homologous recombination; transformation; transgenic plant; ss.
XX
XX      Zea mays.
XX
FH      Key      Location/Qualifiers
FT      CDS      85..1452
FT              /*tag= a
FT              /product= "Maize Ruvb orthologue protein #1"
XX
XX      W020105975-A1.
XX
PD      25-JAN-2001.
XX
PE      13-JUN-2000; 2000WO-US16271.
XX
PR      16-JUL-1999; 99US-0144112.
XX
PA      (PION-) PIONEER HI-BRED INT INC.
XX
P1      Mahajan PB;
XX
WI      WPI: 2001-159537/16.
XX
P-PSDB: AAY72562.
XX
PT      Novel maize Ruvb nucleic acid useful for modulating levels of maize
PT      Ruvb in plants, as probes or amplification primers in the detection,
PT      quantitation or isolation of gene transcripts -
XX
XX      Claim 1; Page 67-69; 87pp; English.
XX
XX      The present sequence is a Zea mays Ruvb orthologue #1 cDNA. Ruvb along
XX      with Ruva catalyses the branch migration process, also known as
XX      heteroduplex extension, in homologous recombination. Ruvb is used for
XX      the control of homologous recombination or transformation efficiency in
XX      transgenic plants. The Ruvb nucleotide may be used as probes or
XX      amplification primers for detecting, quantifying or isolating gene
XX      transcripts, in detecting deficiencies in the mRNA level during gene
XX      transfer for desired transgenic plants, for detecting gene mutations, for
XX      monitoring upregulation of expression or changes in enzyme activity, for
XX      detecting any number of allelic variants, orthologues or paralogues of
XX      the gene, or for site directed mutagenesis in eukaryotic cells. It may
XX      also be used for recombinant expression of its encoded polypeptide, or
XX      for use as immunogen in preparing and/or screening of antibodies, and in
XX      sense or antisense suppression of one or more genes in a host cell,
XX      tissue or plant. The Ruvb proteins may be used in assays to agonise or
XX      antagonise the enzyme function, or as immunogens or antigens for
XX      screening antibodies.
XX
XX      Sequence 1845 BP; 530 A; 378 C; 490 G; 445 T; 2 other:

```

```

Query Match      30.1%; Score 576; DB 22; Length 1845;
Best Local Similarity 98.8%; Pred. No. 2.2e-214;
Matches 1426; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      84 gttggcgcgataagatcagagatgagatgacgtcagctcgaagaagcagcagcagcagc 143
Db      75 gttggcgcgataagatcagagatgagatgacgtcagctcgaagaagcagcagcagcagc 134
QY      144 ccaacccacacataagagactcgcgcctcgcagccatgagatgagatgagatgagatgagat 203
Db      135 ccaacccacacataagagactcgcgcctcgcagccatgagatgagatgagatgagatgagat 194
QY      204 ggggttcgtggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 253
Db      195 ggggttcgtggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 254
QY      264 ccagaagaagaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 323
Db      255 ccagaagaagaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 314
QY      324 ggcgtacgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 383
Db      315 ggcgtacgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 374
QY      384 agatcagaagtgatcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 443
Db      375 agatcagaagtgatcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 434
QY      444 taagctataggttgctgataaagaaacaaagagttatgaaagagagttactga 503
Db      435 taagctataggttgctgataaagaaacaaagagttatgaaagagagttactga 494
QY      504 acttccccaagagagctgaagatcactggtgatatgcaaaaagcattagccatgt 563
Db      495 acttccccaagagagctgaagatcactggtgatatgcaaaaagcattagccatgt 554
QY      564 aatcatcagcttaaaagctgttaaaggaactaagcaacggaagttagattcttaattta 623
Db      555 aatcatcagcttaaaagctgttaaaggaactaagcaacggaagttagattcttaattta 614
QY      624 tgatgctctgatacaagaagaagtgacgtggtgtagatgatatatcatcgaagaactag 683
Db      615 tgatgctctgatacaagaagaagtgacgtggtgtagatgatatatcatcgaagaactag 674
QY      684 tggagcagtgaaagaagtggtgtagatgatatcttctgtacagaatacagatctgaagc 743
Db      675 tggagcagtgaaagaagtggtgtagatgatatcttctgtacagaatacagatctgaagc 734
QY      744 tgaagatgattctcattcccaagaagtgaaagtcataaagaagaagaagaatgtgtcaga 803
Db      735 tgaagatgattgttctcattcccaagaagtgaaagtcataaagaagaagaagaatgtgtcaga 794
QY      804 tctcacacttcattgacctgatacgcagaatgtcagcacacaaggtggtccaaagatttt 863
Db      795 tctcacacttcattgacctgatacgcagaatgtcagcacacaaggtggtccaaagatttt 854
QY      864 gtcccttaaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 923
Db      855 gtcccttaaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 914
QY      924 agaaatataagtgatgataatagatatcagatgaagaagaatttcagagctctgacgttg 983
Db      915 agaaatataagtgatgataatagatatcagatgaagaagaatttcagagctctgacgttg 974
QY      984 tgtttgttcattgataagtgatcacaatgttgatatcgaatgtttcttatacttaacgc 1043
Db      975 tgtttgttcattgataagtgatcacaatgttgatatcgaatgtttcttatacttaacgc 1034
QY      1044 tgcattgagagccattatcacaatcgtgatactcgtcacaataaggggaatgtatgtaa 1103
Db      1035 tgcattgagagccattatcacaatcgtgatactcgtcacaataaggggaatgtatgtaa 1094

```



XX PS Disclosure: Page 77-79; 87pp; English.

CC The present sequence is a *Zea mays* RuVb orthologue #4 cDNA. RuVb along  
 CC with RuVb catalyses the branch migration process, also known as  
 CC heteroduplex extension, in homologous recombination. RuVb is used for  
 CC the control of homologous recombination or transformation efficiency in  
 CC transgenic plants. The RuVb nucleotide may be used as probes or  
 CC amplification primers for detecting, quantifying or isolating gene  
 CC transcripts, in detecting deficiencies in the mRNA level during screening  
 CC for desired transgenic plants, for detecting gene mutations, for  
 CC monitoring upregulation of expression or changes in enzyme activity, for  
 CC detecting any number of allelic variants, orthologues or paralogues of  
 CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
 CC also be used for recombinant expression of its encoded polypeptide, or  
 CC for use as immunogen in preparing and/or screening of antibodies, and in  
 CC sense or antisense suppression of one or more genes in a host cell,  
 CC tissue or plant. The RuVb proteins may be used in assays to agonise or  
 CC antagonise the enzyme function, or as immunogens or antigens for  
 CC screening antibodies.

XX SO Sequence 1898 BP; 532 A; 413 C; 521 G; 432 T; 0 other;

Query Match 4.9%; Score 94; DB 22; Length 1898;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;  
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 gaggtcaagaacacgaggtgctgatagtgaaattccgtagagctatagttgcgtata 465  
 DB 481 gagtcaagaacacgaggtgctgatagtgaaattccgtagagctatagttgcgtata 540  
 QY 466 aaggaacacacagaggttataaggaagaggtta 499  
 DB 541 aaggaacacacagaggttataaggaagaggtta 574

RESULT 6  
 AAT88081  
 ID AAT88081 standard; cDNA; 69 BP.  
 AC AAT88081;  
 AT 14-MAY-1998 (first entry)  
 DE 3' portion of cDNA clone encoding secreted protein AK296.  
 KW Human; secreted protein; research; treatment; AM533;  
 KW 3' portion; ds.  
 OS Homo sapiens.  
 OS WO9739122-A2.  
 PN 23-OCT-1997.  
 PD 11-APR-1997; 97WO-US06042.  
 PF 12-APR-1996; 96US-0631184.  
 PR (MURRO-) MURRO PHARM INC.  
 PA Theoharides TC;  
 PI WPI; 1997-526459/48.  
 DR Human and murine secreted proteins - useful to research or treat  
 PT diseases or disorders related to their function  
 PS Disclosure: Page 88; 140pp; English.  
 CC The present sequence is the 3' portion of a cDNA clone encoding a  
 CC human secreted protein, which may have nutritional uses, or

CC cytokine and cell proliferation/differentiation, immune stimulating  
 CC or suppressing, haematopoiesis regulating, tissue growth,  
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic and  
 CC thrombolytic, receptor/ligand, anti-inflammatory or tumour  
 CC inhibition activities. It can also be used to research or treat  
 CC diseases/disorders related to its function.  
 CC The partial cDNA clone Ap162 was 1st isolated from a human adult  
 CC placenta cDNA library. The partial cDNA clones AK931, AM610, AM340,  
 CC AM82, AK447, AK583, AK533 and AK296 were 1st isolated from a human  
 CC foetal kidney cDNA library. The partial cDNA clones H617 and B89  
 CC were 1st isolated from a human peripheral blood monocyte cell (Th1  
 CC or Th2) cDNA library. The partial cDNA clone AM191 was 1st isolated  
 CC from a human ovary (PA-1 teratocarcinoma) cDNA library. The partial  
 CC cDNA clones AT211, AT205 and AT319 were 1st isolated from a human  
 CC lymphocyte and dendritic cell cDNA library. The partial cDNA clones  
 CC AS34 and AS32 were 1st isolated from a human foetal brain cDNA  
 CC library. The partial cDNA clone AR260 was 1st isolated from a human  
 CC adult retina cDNA library. The partial cDNA clones K640 and K39  
 CC were 1st isolated from a murine bone marrow (stromal cell line  
 CC FCM-4) cDNA library.

XX SO Sequence 69 BP; 48 A; 6 C; 4 G; 9 T; 2 other;

Query Match 1.8%; Score 34; DB 18; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 0.00042;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
 DB 5 taataaaaaaaaaaaaaaaaaaaaaaaaaa 38

RESULT 7  
 AAV02148  
 ID AAV02148 standard; cDNA; 69 BP.  
 AC AAV02148;  
 AT 12-MAY-1998 (first entry)  
 DE Human secreted protein AK296 3' portion including the polyA tail.  
 KW Human; secreted protein; ATCC 98026; cytokine; immunomodulation;  
 KW cell proliferation; differentiation; regulation; ds.  
 OS Homo sapiens.  
 OS WO9739123-A2.  
 PN 23-OCT-1997.  
 PD 14-APR-1997; 97WO-US06139.  
 PF 18-APR-1996; 96US-0634325.  
 PR (GENV) GENETICS INST INC.  
 PA Jacobs K, Lavallic ER, McCoy JM, Merberg D, Racie LA;  
 PI Spaulding V;  
 PI WPI; 1997-526460/48.  
 DR New secreted proteins encoded clones present in ATCC 98026 -  
 PT possibly having cytokine, cell proliferation/differentiation  
 PT regulating, immunomodulating and many other activities  
 PS Disclosure: Page 88; 139pp; English.  
 CC The present sequence encodes a portion of a novel human secreted protein  
 CC deposited under accession number ATCC 98026. The secreted protein can be  
 CC used to determine biological activity, to raise antibodies, as tissue  
 CC markers, to isolate cognate ligands or receptors, to identify agents







PN W09746683-A2.  
 XX 11-DEC-1997.  
 PD  
 XX  
 XX 06-JUN-1997; 97WO-US09878.  
 PR 07-JUN-1996; 96US-0659224.  
 XX  
 XX (GEMV ) GENETICS INST INC.  
 PA  
 XX Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM;  
 PI Merberg D, Racie LA, Spaulding V, Treacy M;  
 DR WPI: 1998-042191/04.  
 XX  
 PT Nucleic acids encoding secreted proteins from clones within ATCC  
 PT 98076 - useful as immuno-modulators, anti-proliferative agents,  
 PT regulators of cell differentiation and tissue growth, etc  
 PS  
 PS Claim 13; Page 57; 99pp; English.  
 XX  
 CC The present sequence represents the nucleotide sequence of the 3'  
 CC portion of clone AZ302.1. The clone was isolated from a human colon  
 CC (Caca-2 adenocarcinoma) cDNA library using probe AAV05749. AZ302.1 is  
 CC a full length clone encoding a secreted protein. AAV05719 provides the  
 CC 5' portion sequence and AAV05718 the internal sequence. The AZ302.1  
 CC sequence shows some homology with an EST sequence. The nucleic acid  
 CC can be used for expression of recombinant proteins, as tissue, molecular  
 CC weight or chromosome markers. Indicators of genetic disorders and  
 CC sources of probes and primers. They can also be used to generate  
 CC anti-protein or anti-DNA antibodies and as components of interaction trap  
 CC assays etc. The protein is useful for raising antibodies, as immunoassay  
 CC reagents and as nutritional supplements. The protein may possibly have  
 CC any of a great variety of therapeutic activities.  
 CC  
 SQ Sequence 85 BP; 80 A; 0 C; 0 G; 5 T; 0 other;

Query Match 1.8%; Score 34; DB 19; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taataaaaaaaaaaaaaaaaaaaaaa 1912  
 ||||||||||||||||||||||||||||  
 DB 5 taataaaaaaaaaaaaaaaaaaaaaa 38

RESULT 13  
 AAQ11760/c  
 ID AAQ11760 standard; DNA; 86 BP.  
 XX  
 AC AAQ11760;  
 XX  
 DT 22-JUL-1991 (first entry)  
 XX  
 DE Self-complementary, T7 promoter hairpin-forming sequence #2.  
 XX  
 KM T7 bacteriophage; DNA-dependent RNA polymerase;  
 KM target sequence amplification; ss.  
 KW  
 XX  
 OS Synthetic.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT 1..86  
 FT stem\_loop /tag= a  
 FT /note= "number of T's in loop is 50"  
 PN BP427073-A.  
 XX  
 PD 15-MAY-1991.  
 XX  
 PF 27-OCT-1990; 90EP-0120650.  
 XX

PR 23-AUG-1990; 90US-0569991.  
 PR 09-NOV-1989; 89US-0434372.  
 XX  
 XX (MOLE-) MOLECULAR DIAGNOSTI.  
 PA  
 XX  
 XX Dattagupta N;  
 PI  
 DR WPI: 1991-141638/20.  
 XX  
 PT Nucleic acid probe for amplification and detection of target  
 PT sequence - capable of forming ligatable hairpin structured  
 PT promoter and transcribing target sequence, is sensitive and  
 PT useful in medical diagnosis  
 PS  
 PS Claim 2; Page 11; 15pp; English.  
 XX  
 CC The sequence is an example of a preferred T7 RNA polymerase  
 CC promoter for use in the invention. There can be from 2 to 50 T's in  
 CC the loop region. A probe sequence is ligated to the 3' end of the  
 CC promoter region. Upon hybridisation of the probe to a target sequence  
 CC and ligation of the hybridised target sequence to the 5' end of the  
 CC hairpin-forming sequence, the target sequence can be transcribed by  
 CC T7 RNA polymerase. This allows the target sequence to be amplified  
 CC (and detected) using a single oligonucleotide component (c.f. PCR).  
 CC See also AAQ11759 and AAQ11761-Q11764.  
 XX  
 SQ Sequence 86 BP; 12 A; 6 C; 6 G; 62 T; 0 other;

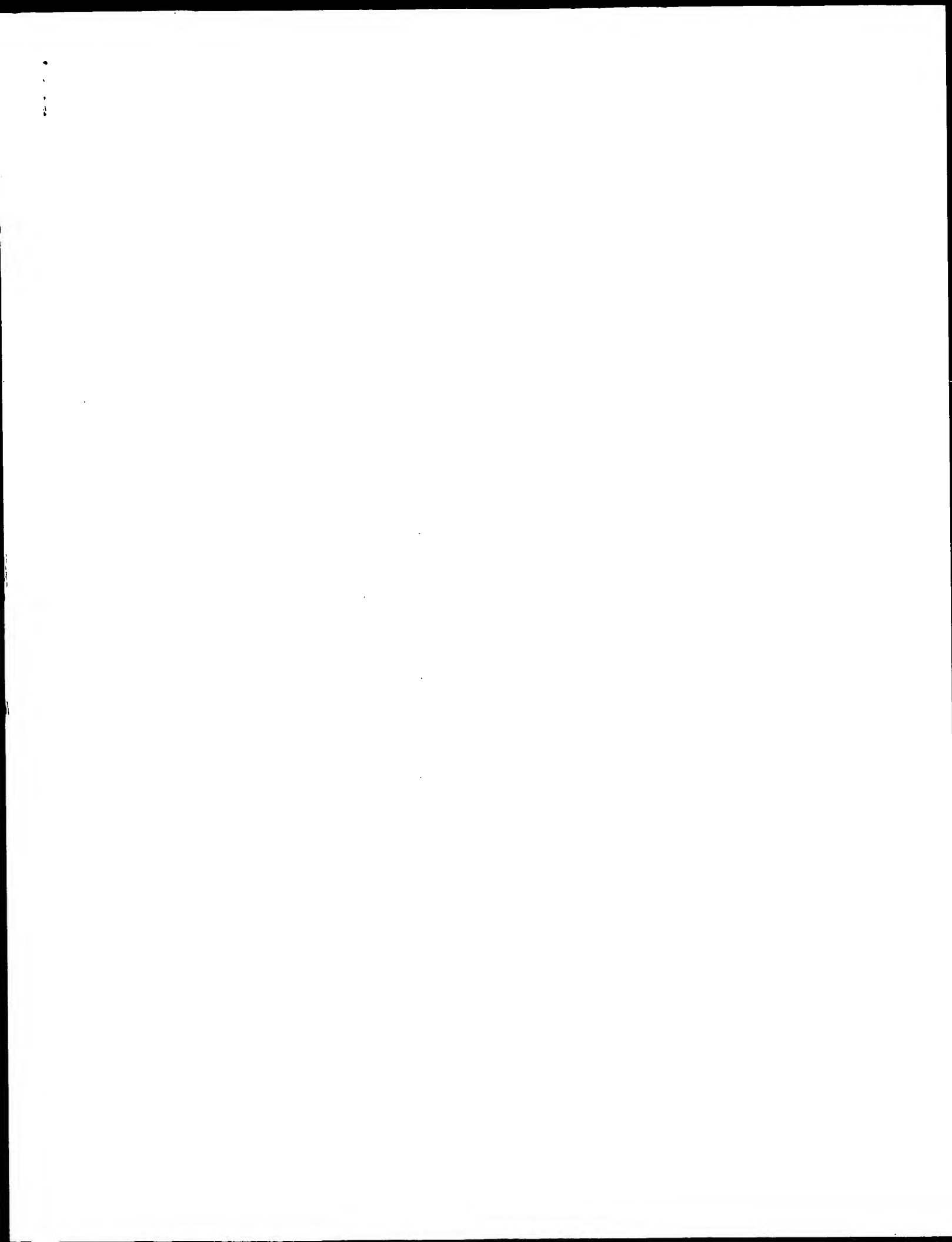
Query Match 1.8%; Score 34; DB 12; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 0.0004;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taataaaaaaaaaaaaaaaaaaaaaa 1912  
 ||||||||||||||||||||||||||||  
 DB 70 TAAAAAAAAAAAAAAAAAAAAAAAAAAAA 37

RESULT 14  
 AAC78380  
 ID AAC78380 standard; CDNA; 89 BP.  
 XX  
 AC AAC78380;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated gene sequence SEQ ID NO:774.  
 XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;  
 KW antidiabetic; antidiabetic; antirheumatic; antiarthritic; antiviral;  
 KW antiinflammatory; antihypertensive; antiallergic; antibacterial; antiandrogenic;  
 KW dermatologic; neuroprotective; thrombolytic; coagulant; motropic;  
 KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening; ss.  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 PN WO20005350-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05882.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 08:43:39 : Search time 100.16 Seconds  
(without alignments)  
4323.342 Million cell updates/sec

Title: US-09-589-510-3  
Perfect score: 1912  
Sequence: 1 accacagctccgcaattt.....aaaaaaaaaaaaaaaa 1912

Scoring table:  
OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 351203 segs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :  
1: /cgn2-6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2-6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2-6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2-6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2-6/ptodata/2/ina/PTUS.COMB.seq.\*  
6: /cgn2-6/ptodata/2/ina/Backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	1.8	55	1	US-08-113-646A-41 Sequence 41, Appl
2	34	1.8	114	1	US-08-120-827-99 Sequence 99, Appl
3	34	1.8	114	1	US-08-478-675-99 Sequence 99, Appl
4	34	1.8	120	1	US-08-153-051B-28 Sequence 28, Appl
5	34	1.8	120	1	US-08-060-952C-44 Sequence 44, Appl
6	34	1.8	120	2	US-08-151-477A-28 Sequence 28, Appl
7	34	1.8	120	3	US-08-819-867-58 Sequence 58, Appl
8	34	1.8	140	1	US-08-628-417-5 Sequence 5, Appl
9	34	1.8	141	3	US-08-737-078A-1 Sequence 1, Appl
10	34	1.8	141	5	PCF-US94-04706-1 Sequence 1, Appl
11	34	1.8	144	1	US-08-702-344-26 Sequence 26, Appl
12	34	1.8	216	1	US-08-686-878A-34 Sequence 15, Appl
13	34	1.8	222	4	US-08-481-190-15 Sequence 15, Appl
14	34	1.8	222	5	PCF-US93-00869-15 Sequence 6, Appl
15	34	1.8	240	1	US-08-628-417-6 Sequence 3, Appl
16	34	1.8	289	2	US-08-341-568-3 Sequence 3, Appl
17	34	1.8	289	2	US-08-911-020-3 Sequence 27, Appl
18	34	1.8	340	1	US-08-171-385-27 Sequence 27, Appl
19	34	1.8	340	3	US-08-361-441B-27 Sequence 2, Appl
20	34	1.8	347	1	US-08-104-072B-2 Sequence 14, Appl
21	34	1.8	350	1	US-08-171-385-11 Sequence 11, Appl
22	34	1.8	350	1	US-08-248-016-11 Sequence 11, Appl
23	34	1.8	350	1	US-08-451-501-11 Sequence 14, Appl
24	34	1.8	350	3	US-08-361-441B-14 Sequence 11, Appl
25	34	1.8	350	5	PCF-US95-06761-11 Sequence 39, Appl
26	34	1.8	361	2	US-08-465-380-39 Sequence 39, Appl
27	34	1.8	361	2	US-08-486-397-39 Sequence 39, Appl

28	34	1.8	361	2	US-08-486-399-39 Sequence 39, Appl
29	34	1.8	361	2	US-08-461-965-39 Sequence 39, Appl
30	34	1.8	361	2	US-08-634-641-39 Sequence 39, Appl
31	34	1.8	361	3	US-09-249-471-39 Sequence 39, Appl
32	34	1.8	361	3	US-09-249-472-39 Sequence 39, Appl
33	34	1.8	361	3	US-09-249-451-39 Sequence 39, Appl
34	34	1.8	361	3	US-08-809-455-39 Sequence 39, Appl
35	34	1.8	361	3	US-09-249-461-39 Sequence 39, Appl
36	34	1.8	361	3	US-09-249-448-39 Sequence 23, Appl
37	34	1.8	375	3	US-08-946-026-23 Sequence 13, Appl
38	34	1.8	399	1	US-07-885-970A-13 Sequence 12, Appl
39	34	1.8	399	1	US-08-530-797-12 Sequence 12, Appl
40	34	1.8	399	1	US-08-298-829-13 Sequence 12, Appl
41	34	1.8	399	1	US-08-787-835-12 Sequence 12, Appl
42	34	1.8	399	2	US-08-688-988-44 Sequence 44, Appl
43	34	1.8	444	3	US-08-524-757-1 Sequence 1, Appl
44	34	1.8	458	1	US-08-924-759-7 Sequence 7, Appl
45	34	1.8	458	2	US-08-924-759-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-08-113-646A-41  
Sequence 41, Application US/08113646A  
Patent No. 5578468  
GENERAL INFORMATION:  
APPLICANT: PICKUP, David J.  
APPLICANT: PATEL, Dhavalakumar  
APPLICANT: ANTICZAK, James B.  
TITLE OF INVENTION: SITE-SPECIFIC RNA CLEAVAGE  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHIE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/113.646A  
FILING DATE: 31-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/084,406  
FILING DATE: 10-AUG-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-20  
TELEPHONE/COMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: RNA (genomic)  
US-08-113-646A-41  
Query Match 1.8%; Score 34; DB 1; Length 55;  
Best Local Similarity 97.1%; Pred. No. 16-05; Indels 0; Gaps 0;  
Matches 33; Conservative 1; Mismatches 0;

OY 1879 taataaaataaaataaaataaaataaa 1912  
:|||||  
DB 15 UAAAAAAAAAAAAAAAAAAAAAAAAAAAA 48

## RESULT 2

US-08-120-827-99  
; Sequence 99, Application US/08120827  
; Patent No. 5525495  
; GENERAL INFORMATION:  
; APPLICANT: KEENE, JACK D.  
; APPLICANT: KING, PETER H.  
; APPLICANT: LEVINE, TODD  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE  
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS  
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/120,827  
; FILING DATE: 15-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5525495man F.  
; REGISTRATION NUMBER: 24,618  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
US-08-120-827-99

Query Match 1.8%; Score 34; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 9.7e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taataaaataaaataaaataaaataaa 1912  
:|||||  
DB 80 TAAAAAAAAAAAAAAAAAAAAAAAAAAAA 113

## RESULT 3

US-08-478-675-99  
; Sequence 99, Application US/08478675  
; Patent No. 5773246  
; GENERAL INFORMATION:  
; APPLICANT: KEENE, JACK D.  
; APPLICANT: KING, PETER H.  
; APPLICANT: LEVINE, TODD  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE  
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF  
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,675  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/120,827  
; FILING DATE: 15-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5773246man F.  
; REGISTRATION NUMBER: 24,618  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
US-08-478-675-99

NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,675  
FILING DATE: 07-JUN-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/120,827  
FILING DATE: 15-SEP-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5773246man F.  
REGISTRATION NUMBER: 24,618  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)413-3000  
TELEFAX: (703)413-2220

## INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:  
LENGTH: 114 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
US-08-478-675-99

Query Match 1.8%; Score 34; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 9.7e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taataaaataaaataaaataaaataaa 1912  
:|||||  
DB 80 TAAAAAAAAAAAAAAAAAAAAAAAAAAAA 113

## RESULT 4

US-08-153-051B-28  
; Sequence 28, Application US/08153051B  
; Patent No. 5645986  
; GENERAL INFORMATION:  
; APPLICANT: Michael D. West  
; APPLICANT: Jerry W. Shay  
; APPLICANT: Woodring E. Wright  
; APPLICANT: Elizabeth Blackburn  
; APPLICANT: Nam Woo Kim  
; APPLICANT: Calvin B. Harley  
; APPLICANT: Scott L. Weinrich  
; APPLICANT: Catherine Strahl  
; APPLICANT: Michael J. McEachern  
; TITLE OF INVENTION: Homayoun Vaziri  
; TITLE OF INVENTION: THERAPY RELATED TO TELEOMERE  
; TITLE OF INVENTION: CONDITIONS RELATED TO TELEOMERE  
; TITLE OF INVENTION: LENGTH AND/OR TELEOMERASE ACTIVITY  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California



COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSEO Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,051B  
FILING DATE: No. 564596ember 12, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/038,766  
FILING DATE: March 24, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 204/195  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-153-051B-28

Query Match 1.8%; Score 34; DB 1; Length 120;  
Best Local Similarity 100.0%; Pred. No. 9.7e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
DB 73 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 106

RESULT 5  
US-08-060-952C-44  
Sequence 44, Application US/08060952C  
Patent No. 5695932  
GENERAL INFORMATION:  
APPLICANT: Michael D. West  
APPLICANT: Jerry W. Shay  
APPLICANT: Woodring E. Wright  
APPLICANT: Elizabeth Blackburn  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF CONDITIONS  
RELATED TO TELOMERE LENGTH AND/OR  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/060,952C  
FILING DATE: May 13, 1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/882,438

FILING DATE: May 13, 1992  
APPLICATION NUMBER: 08/038,766  
FILING DATE: March 24, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 202/045  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-060-952C-44

Query Match 1.8%; Score 34; DB 1; Length 120;  
Best Local Similarity 100.0%; Pred. No. 9.7e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1879 taaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
DB 73 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 106

RESULT 6  
US-08-151-477A-28  
Sequence 28, Application US/08151477A  
Patent No. 5830644  
GENERAL INFORMATION:  
APPLICANT: Michael D. West  
APPLICANT: Jerry W. Shay  
APPLICANT: Woodring E. Wright  
APPLICANT: Elizabeth Blackburn  
APPLICANT: Nam Woo Kim  
APPLICANT: Calvin B. Harley  
APPLICANT: Scott L. Weinrich  
APPLICANT: Catherine Strahl  
APPLICANT: Michael J. McEachern  
APPLICANT: Homayoun Vaziri  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF  
CONDITIONS RELATED TO TELOMERE  
LENGTH AND/OR TELOMERASE ACTIVITY  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSEO Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/151,477A  
FILING DATE: No. 5830644ember 12, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/038,766  
FILING DATE: March 24, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 202/189  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-151-477A-28

Query Match  
Best Local Similarity 1.8%; Score 34; DB 2; Length 120;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
Db 73 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 106

RESULT 7  
US-08-819-867-58

; Sequence 58, Application US/08819867  
; Patent No. 6007989

; GENERAL INFORMATION:

; APPLICANT: Michael D. West  
; APPLICANT: Calvin B. Harley  
; APPLICANT: Scott L. Weinrich  
; APPLICANT: Catherine M. Strahl  
; APPLICANT: Michael J. McEachern  
; APPLICANT: Jerry Shay  
; APPLICANT: Woodring E. Wright  
; APPLICANT: Elizabeth H. Blackburn  
; APPLICANT: Nam Woo Kim  
; APPLICANT: Homayoun Vaziri  
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: CONDITIONS RELATED TO  
; TITLE OF INVENTION: TELOMERE LENGTH AND/OR  
; TITLE OF INVENTION: TELOMERASE ACTIVITY  
; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; MEDIUM TYPE: storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/819,867  
; FILING DATE: March 14, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/153,051  
; FILING DATE: No. 6007989ember 12, 1993  
; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Chambers, Daniel M.

; REGISTRATION NUMBER: 34,561

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:  
LENGTH: 120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-819-867-58

Query Match  
Best Local Similarity 1.8%; Score 34; DB 3; Length 120;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
Db 73 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 106

RESULT 8

US-08-628-417-5

; Sequence 5, Application US/08628417

; Patent No. 5627054

; GENERAL INFORMATION:

; APPLICANT: GILLESPIE, DAVID  
; TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC  
; TITLE OF INVENTION: POLYMERASE CHAIN REACTION  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL  
; ADDRESS: DEFENSE COMMAND  
; STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)  
; CITY: ABERDEEN PROVING GROUND  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 21010-5423

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,417  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIFONT, ULYSSES J  
; REGISTRATION NUMBER: 39,908  
; REFERENCE/DOCKET NUMBER: DAM 398-94  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 410-671-1158  
; TELEFAX: 410-671-2534  
; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 140 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: oligodeoxynucleotide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: YES  
US-08-628-417-5

Query Match  
Best Local Similarity 1.8%; Score 34; DB 1; Length 140;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
Db 39 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 72

RESULT 9

US-08-737-078A-1/c

```
; Sequence 1, Application US/08737078A
; Patent No. 6027934
; GENERAL INFORMATION:
; APPLICANT: Powell, Curtis
; TITLE OF INVENTION: VACCINE FOR, DIAGNOSTIC ASSAY FOR AND
; TITLE OF INVENTION: METHOD OF TREATING PARASITIC HEMOFLAGELLATE PROTOZOA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.5
; SOFTWARE: WordPerfect 3.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,078A
; FILING DATE: 16-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04931
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,464
; FILING DATE: 13-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/365,413
; FILING DATE: 13-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: POWELL 203-PCT-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-737-078A-1

Query Match 1.8%; Score 34; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataaaaaaaaaaaaaaaaaaaaaa 1912
DB 61 TAAAAAaaaaaaaaaaaaaaaaaaaaa 28

RESULT 10
PCT-US94-04706-1/C
; Sequence 1, Application PC/TUS9404706
; GENERAL INFORMATION:
; APPLICANT: Powell, Curtis
; TITLE OF INVENTION: VACCINE FOR, DIAGNOSTIC ASSAY FOR AND
; TITLE OF INVENTION: METHOD OF TREATING PARASITIC
; TITLE OF INVENTION: HEMOFLAGELLATE PROTOZOA
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Horn Kramer & Woods
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04706
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04931
; FILING DATE: 11-JUN-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,464
; FILING DATE: 13-JUN-91
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/365,413
; FILING DATE: 13-JUN-89
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: POWELL 203-PCT-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US94-04706-1
```

```
Query Match 1.8%; Score 34; DB 5; Length 141;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataaaaaaaaaaaaaaaaaaaaaa 1912
DB 61 TAAAAAaaaaaaaaaaaaaaaaaaaaa 28

RESULT 11
US-08-702-344-26
; Sequence 26, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavalie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344
; FILING DATE:
```

CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 144 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-702-344-26

Query Match 1.8%; Score 34; DB 1; Length 144;  
Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataataataataataataataataataata 1912  
|||||  
Db 22 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55

RESULT 12  
US-08-686-878A-34  
Sequence 34, Application US/08686878A  
Patent No. 5708157  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John  
APPLICANT: Lavallie, Edward  
APPLICANT: Racie, Lisa  
APPLICANT: Merberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Evans, Cheryl  
APPLICANT: Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
TITLE OF INVENTION: ENCODING THEM  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/686,878A  
APPLICATION NUMBER: US/08/686,878A  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 216 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-686-878A-34

Query Match 1.8%; Score 34; DB 1; Length 216;  
Best Local Similarity 100.0%; Pred. No. 9.3e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataataataataataataataataataata 1912  
|||||  
Db 168 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 201

RESULT 13  
US-08-481-190-15  
Sequence 15, Application US/08481190  
Patent No. 6160204  
GENERAL INFORMATION:  
APPLICANT: John C. Steffens  
TITLE OF INVENTION: Polyphehol Oxidase CDNA  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Yahwak & Associates  
STREET: 25 Skytop Drive  
CITY: Trumbull  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06611  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,190  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 203,533  
FILING DATE: 02-24-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: George M. Yahwak  
REGISTRATION NUMBER: 26,824  
REFERENCE/DOCKET NUMBER: UA 816 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203)268-1951  
TELEFAX: (203)268-1951  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-481-190-15

Query Match 1.8%; Score 34; DB 4; Length 222;  
Best Local Similarity 100.0%; Pred. No. 9.3e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 taataataataataataataataataataata 1912  
|||||  
Db 176 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 209

RESULT 14  
PCT-US93-00869-15  
Sequence 15, Application PC/TUS9300869  
GENERAL INFORMATION:  
APPLICANT: John C. Steffens  
TITLE OF INVENTION: Polyphehol Oxidase CDNA: Cloning  
TITLE OF INVENTION: and Applications  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Yahwak & Associates

STREET: 25 Skytop Drive  
CITY: Trumbull  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06611  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00869  
FILING DATE: 19930129  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: George M. Yahwak  
REGISTRATION NUMBER: 26,824  
REFERENCE/DOCKET NUMBER: CRF D-1057  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203)268-1951  
TELEFAX: (203)268-1951  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
PCT-US93-00869-15

Query Match 1.8%; Score 34; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 9.3e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
Db 176 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 209

RESULT 15  
US-08-628-417-6  
Sequence 6, Application US/08628417  
Patent No. 5627054  
GENERAL INFORMATION:  
APPLICANT: GILLESPIE, DAVID  
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL  
ADDRESSEE: DEFENSE COMMAND  
STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)  
CITY: ABERDEEN PROVING GROUND  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21010-5423  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,417  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BIFFONI, ULYSSES J  
REGISTRATION NUMBER: 39,908  
REFERENCE/DOCKET NUMBER: DAM 398-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 410-671-1158  
TELEFAX: 410-671-2534

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: oligodeoxynucleotide  
HYPOTHEICAL: NO  
ANTI-SENSE: YES  
US-08-628-417-6

Query Match 1.8%; Score 34; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 9.2e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1879 taataaaaaaaaaaaaaaaaaaaaaaaaaa 1912  
|||||  
Db 39 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 72

Search completed: November 13, 2001, 09:53:18  
Job time: 4179 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 07:53:44 ; Search time 1462.74 Seconds  
(without alignments)  
12356.164 Million cell updates/sec

Title: US-09-589-510-3  
Perfect score: 1912  
Sequence: 1 acccagcgctccgaattt.....aaaaaaaaaaaaaaaaaaaaa 1912

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est33:\*  
26: gb\_est34:\*  
27: gb\_est35:\*  
28: gb\_est36:\*  
29: gb\_est37:\*  
30: gb\_est38:\*  
31: gb\_est39:\*  
32: gb\_est40:\*  
33: em\_estba:\*  
34: em\_estfun:\*  
35: em\_esthum1:\*  
36: em\_esthum2:\*  
37: em\_esthum3:\*  
38: em\_esthum4:\*  
39: em\_esthum5:\*  
40: em\_esthum6:\*  
41: em\_esthum7:\*  
42: em\_esthum8:\*  
43: em\_esthum9:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estin5:\*  
68: em\_estom1:\*  
69: em\_estom2:\*  
70: em\_estov1:\*  
71: em\_estov2:\*  
72: em\_estp11:\*  
73: em\_estp12:\*  
74: em\_estp13:\*  
75: em\_estp14:\*  
76: em\_estp15:\*  
77: em\_estp16:\*  
78: em\_estp17:\*  
79: em\_estp18:\*  
80: em\_estp19:\*  
81: em\_estp110:\*  
82: em\_estro1:\*  
83: em\_estro2:\*  
84: em\_estro3:\*  
85: em\_estro4:\*  
86: em\_estro5:\*  
87: em\_estro6:\*  
88: em\_estro7:\*  
89: em\_estro8:\*  
90: em\_estro9:\*  
91: em\_estro10:\*  
92: em\_estro11:\*  
93: em\_estro12:\*  
94: em\_estro13:\*  
95: em\_estro14:\*  
96: em\_estro15:\*  
97: em\_estro16:\*  
98: em\_estro17:\*  
99: em\_estro18:\*  
100: em\_estro19:\*  
101: em\_estro20:\*  
102: gb\_est25:\*  
103: gb\_est26:\*  
104: gb\_est27:\*  
105: gb\_est28:\*  
106: gb\_est29:\*  
107: gb\_est30:\*  
108: gb\_est31:\*  
109: gb\_est32:\*  
110: gb\_est41:\*  
111: gb\_est42:\*  
112: gb\_est43:\*  
113: gb\_est44:\*  
114: gb\_est45:\*  
115: gb\_est46:\*  
116: gb\_est47:\*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

---

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.







QY	579	gacgtttaaaggagactaagcaactgaaagttagatcttcattcaattatgatgctcgatcaa	638
Db	181	GACGTTTAAAGGAGCTAAGCAACTGAAGTAGATTCTTCAATTATATGATGCTGATCA	240
QY	639	ggaagaag 646	
Db	241	GGAAGAAG 248	
RESULT	4		
BE512425/c			
LOCUS	BE512425	549 bp	mRNA
DEFINITION	946071A06.x1 946	- tassell primordium prepared by Schmidt	lab Zea
ACCESSION	BE512425		
VERSION	BE512425.1	GI:9733673	
KEYWORDS	EST		
SOURCE			
ORGANISM			
	Zea mays.		
	Zea mays.		
	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC		
	clade; Panicoidae; Andropogoneae; Zea.		
	1 (bases 1 to 549)		
REFERENCE			
AUTHORS	Walbot,V.		
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford		
JOURNAL	University		
COMMENT	Unpublished (1999)		
	Contact: Walbot V		
	Department of Biological Sciences		
	Stanford University		
	855 California Ave, Palo Alto, CA 94304, USA		
	Tel: 650 723 2227		
	Fax: 650 725 8221		
	Email: walbot@stanford.edu		
	Plate: 946071 row: A column: 06.		
FEATURES			
Source			
	location/Qualifiers		
	1..549		
	/organism="Zea mays"		
	/cultivar="OH43"		
	/db_xref="taxon:4577"		
	/clone_lib="946 - tassell primordium prepared by Schmidt		
	lab"		
	/russue_type="tassels"		
	/dev_stage="just after the transition from vegetative to		
	inflorescence development"		
	/lab_host="X10LR"		
	/note="Organ: tassels; Vector: HybriZAP; Site1: EcoRI;		
	Site2: XhoI; George Chuck dissected immature tassels		
	between 1mm and 3mm. Sharon Stanfield prepared the cDNA		
	library in HybriZAP. Sample insert size range was 350 bp		
	to 3 Kb with a 1 Kb average."		
	150 a 146 c 116 g 137 t		
BASE COUNT			
ORIGIN			
Query Match	7.0%;	Score 134;	DB 136; Length 549;
Best Local Similarity	99.0%;	Pred. NO. 8.4e-57;	
Matches 284; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
QY	1240	attgatatgatatgaagaaagttctgcttatttaagcgcgagatctgcgacgacagacatcttg	1299
Db	519	ATTGATATGGAAGAAAGCTTCTTATTATTAAGCGAGATGGACAGCAGACATCTTTTA	460
QY	1300	agacatgcatcattcaattgatatcacctgcgcagcgtggtcacaagaactatggaagag	1359
Db	459	AGACATGCTATTCAATTGATATACCTGCGCAGCTGTCTTCAAAAGACTAATGGAAGAG	400
QY	1360	aaaatctgcaaggtcgatctcgaaggaagtcaagtgggctctatttggatgccaactctcg	1419
Db	399	AAAAATATGCAAGCGTCGATCTCGAGGAAGTCAGTGGGCTCTATTGGATGCCAAATCTCG	340
QY	1420	gctgcgcgtctccagagagcaacaagaatatcatcacttaagtatttgatcttcctgctg	1479

Db	339	GCTGGGCGTCCACAGGACCACAAGAATACATCACCTTGATTGGATCACTGTGCT	280
QY	1480	ggaagtcctcgaagaagtgttagttgcacagctcgaaatcatctagt	1526
Db	279	GGAAGTCCTGAAAGAAATGTAGTTGCCAGCTCGAAGATCATCTAGT	233
RESULT	5		
LOCUS	BE640273/c		
DEFINITION	946082B06.xl 946 - tassal primordium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.	EST	30-AUG-2000
ACCESSION	BE640273		
VERSION	BE640273.1	GI:9953690	
KEYWORDS	EST.		
SOURCE	Zea mays.		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 375)		
AUTHORS	Walbot,V.		
TITLE	Maze ESTs from various cDNA libraries sequenced at Stanford University		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946082 row: B column: 06. location/Qualifiers		
FEATURES	source		
	1..375	/organism="Zea mays"	
	/cultivar="OH43"		
	/db_xref="taxon:4577"		
	/clone_lhb="946 - tassal primordium prepared by Schmidt lab"		
	/tissue_type="tassels"		
	/dev_stage="just after the transition from vegetative to inflorescence development"		
	/lab_host="XLOLR"		
	/note="Organ: tassels; Vector: HybridAP, Site_1: EcorI; Site_2: XhoI; George Chuck dissected immature tassels between Imm and Imm. Sample Stanfield prepared the cDNA library in HybridAP. Shanon insert size range was 350 bp to 3 kb with a 1 kb average."		
BASE COUNT	106 a 106 c 79 g 84 t		
ORIGIN			
Query Match	5.8%; Score 110; DB 138; Length 375;		
Best Local Similarity	99.4%; Pred. No.1.3e-44;		
Matches 160; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	1366	tgcaggctgatctcgagaagtcagtgggctctatttgatgccaaactccctcgctcgg	1425
DB	364	TCGCAAGCGCTGATCTCGAGGANAFTCAGTGGGCTCTATTGGATGCCAAATCTCGGCTCGG	305
QY	1426	ctgctccaggagcaacaagaatacacatcacctagattgatatctctgtcgtygaagt	1485
DB	304	CTGCTCCAGAGACCAACAAAGATAACATCACCTGATTTGGATCACCTGTGTGGAAGT	245
QY	1486	ctcgaagaagaatgttagttgcacagctcgaaatcatctagt	1526
DB	244	CTCGAAGAGAAATGTAGTTGCCAGCTCGAAGATCATCTAGTG	204
RESULT	6		
LOCUS	BG316941/C		
	BG316941	109 bp	mRNA
			EST
			26-FEB-2001

DEFINITION 947024C05.x2 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA sequence.  
 ACCESSION BG316941  
 VERSION BG316941.1 GI:13126371  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 109)  
 AUTHORS Walbot, V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 947024 row: C column: 05.  
 FEATURES  
 source location/Qualifiers  
 1..109  
 /organism="Zea mays"  
 /cultivar="W23"  
 /db\_xref="taxon:4577"  
 /clone\_lib="947 - 2 week shoot from Barkan lab"  
 /tissue\_type="leaf and stem, including leaf base"  
 /dev\_stage="2 week old seedling (3 leaves)"  
 /lab\_host="X11-Blue"  
 /note="Organ: shoot; Vector: lambda ZAP (pBluescript SK-); Site 1: EcoRI; Site 2: XhoI; directionally cloned using StrataGene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10<sup>5</sup> independent recombinant phage. The plants were greenhouse grown."  
 BASE COUNT 30 a 29 c 17 g 33 t  
 ORIGIN  
 Query Match 5.7%; Score 109; DB 152; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-44;  
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1214 tggcgtccgagcacaagtggagagattgatgatgaagaagcttctatttag 1273  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 109 tggcgtatccgagacacaaagtgagagattgatgaagaagcttctatttag 50  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1274 ggcagatcgagacagacatctttagagacatgcatcattgattatc 1322  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 49 cgcagatcgagacagacatctttagagacatgcatcattgattatc 1  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 7  
 LOCUS AW927557 512 bp mRNA EST 30-MAY-2000  
 DEFINITION 945018G05.X2 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.  
 ACCESSION AW927557  
 VERSION AW927557.1 GI:8102879  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 512)  
 AUTHORS Walbot, V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
 JOURNAL Unpublished (1999)

COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 945018 row: G column: 05.  
 FEATURES  
 source location/Qualifiers  
 1..512  
 /organism="Zea mays"  
 /cultivar="W23"  
 /db\_xref="taxon:4577"  
 /clone\_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"  
 /tissue\_type="tassel, kernel, silk, husk, root, leaf"  
 /dev\_stage="fully-grown"  
 /lab\_host="DH10B"  
 /note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site 1: EcoRI; cDNA library from fully differentiated maize tissues from an active mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."  
 BASE COUNT 148 a 98 c 151 g 115 t  
 ORIGIN  
 Query Match 4.9%; Score 94; DB 122; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-36;  
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 406 gaggtcaagaagaactggggtgctatgaaatttcgtagacttaagtgtgctata 465  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 177 GAGGTCAAGAAAGCTGGGtGCTGATGGAATTTCCGTAGACTTATGAGTTTGCCTATA 236  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 466 aaggaacaagaagggtttatgaaggaaggtta 499  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 237 AAGGAAACAAAGAGTGTATGAAAGGAGAGCTTA 270  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 8  
 LOCUS AW927556 555 bp mRNA EST 30-MAY-2000  
 DEFINITION 945018G05.X1 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.  
 ACCESSION AW927556  
 VERSION AW927556.1 GI:8102878  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 555)  
 AUTHORS Walbot, V.  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 945018 row: G column: 05.  
 FEATURES  
 source location/Qualifiers  
 1..555  
 /organism="Zea mays"  
 /cultivar="W23"  
 /db\_xref="taxon:4577"  
 /clone\_lib="945 - Mixed adult tissues from Walbot lab,

same as 707 (SK)"  
/tissue\_type="tassel", kernel, silk, husk, root, leaf"  
/dev\_stage="fully-grown"  
/lab\_host="PH10B"  
/note="Organ: tassel, kernel, silk, husk, root, leaf;  
vector: PGAD10; Site\_1: Ecor1; cDNA library from fully  
differentiated maize tissues from an active Mutator plant.  
tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,  
root, leaf). Unidirectionally cloned. New library number  
given to library 707 for additional sequencing."  
BASE COUNT 156 a 111 c 165 g 123 t  
ORIGIN

Query Match 4.9%; Score 94; DB 122; Length 555;  
Best Local Similarity 100.0%; Pred. No. 1.8e-36;  
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 gaagtcagaagaactgagtcgtcatggaataatccgtagagctatagttgcgtata 465  
|||||  
Db 194 GAGGTCAAGAAAGCTGAGGTGCTGATGGAATAATTCCTAGACTATAGTTTGCCTATA 253  
|||||  
QY 466 aaggaacacaaagagtttatgaaggaaggtta 499  
|||||  
Db 254 AAGGAAACAAAGAGGTTTATGAAGAGAGGTTA 287  
|||||

RESULT 9  
BG316942 169 bp mRNA EST 26-FEB-2001  
LOCUS 947024C05.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA  
DEFINITION  
sequence.  
ACCESSION BG316942  
VERSION BG316942.1 GI:13126372  
KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 169)  
Walbot.V.

REFERENCE  
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford  
TITLE University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 947024 row: C column: 05.

FEATURES  
source Location/Qualifiers  
1. 169

/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="947 2 week shoot from Barkan lab"  
/tissue\_type="leaf and stem, including leaf base"  
/dev\_stage="2 week old seedling (3 leaves)"  
/lab\_host="XL1-Blue"  
/note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);  
Site\_1: Ecor1; Site\_2: Xho1; Directionally cloned using  
Stratagene's Unizap XR cDNA cloning kit with the 5' end  
at the Ecor1 site. The library represents 8 x 10e5  
independent recombinant phage. The plants were greenhouse  
grown."  
BASE COUNT 56 a 27 c 36 g 50 t  
ORIGIN

Query Match

4.6%; Score 88; DB 152; Length 169;

Best Local Similarity 100.0%; Pred. No. 2.1e-33;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 883 atgaacacgaagaactcgaatccggaataacacgacccgaagaataaagtgta 942  
|||||  
Db 1 ATGAACACGACGAAAGCTGAATACCGAAACACTACCCCAAGAAATTAAGTGCTA 60  
|||||  
QY 943 aatagatatcatgtaggaattgcag 970  
|||||  
Db 61 AATAGATATATCATGATGAGGAATTCAG 88  
|||||

RESULT 10  
BG464307 535 bp mRNA EST 20-MAR-2001  
LOCUS EM1\_72.B07.b1.A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA  
DEFINITION  
sequence.  
ACCESSION BG464307  
VERSION BG464307.1 GI:13392679  
KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 535)  
Reid S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.  
An EST database from Sorghum: developing embryos  
Unpublished (2000)  
CONTACT: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu

REFERENCE  
AUTHORS Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
TITLE Seq primer: JEN REV  
JOURNAL High quality sequence stop: 517  
COMMENT POLYA=NO.

FEATURES  
source Location/Qualifiers  
1. 535  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Embryo 1 (EM1)"  
/note="Organ: Embryos germinated for 24 hr; Vector:  
pBluescript II from Lambda Zap II; Site\_1: Xho1; Site\_2:  
Ecor1; The library was made from poly-A RNA in the cloning  
vector lambda ZAP II. Clones to be sequenced were  
prepared by mass excision."  
BASE COUNT 127 a 126 c 183 g 99 t  
ORIGIN

Query Match 4.3%; Score 82; DB 154; Length 535;  
Best Local Similarity 99.2%; Pred. NO. 2.2e-30;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 tagccagaagctcgagcgaagtcctctctgctcatgtagtagatcagaagtgta 400  
|||||  
Db 293 TAGCCAGAGACTGCGCGACAGAGTCCCTTCTGTCTATGATGATCAAGATGTA 352  
|||||  
QY 401 cctcgagagtcagaagaactcgaagtcgtatgtaaaatccgtagagctataggttc 460  
|||||  
Db 353 CCTCAGAGTCAAGAAAGCTGAGGTGCTGATGGAATAATTCCTAGAGCTATAGTTGC 412  
|||||

QY 461 gtataaagaaaa 473  
|||||  
Db 413 GTATTAAGGAAAA 425  
|||||

RESULT 11  
BE512214/c 519 bp mRNA EST 07-AUG-2000  
LOCUS 94606A01.x1 946 - tassal primordium prepared by Schmidt lab Zea  
DEFINITION mays cDNA, mRNA sequence.  
ACCESSION BE512214  
VERSION BE512214.1 GI:9733357  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade: Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 519)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946066 row: A column: 01.

FEATURES  
source  
1..519  
/organism="Zea mays"  
/cultivar="OH43"  
/db\_xref="taxon:4577"  
/clone\_lib="946 - tassal primordium prepared by Schmidt  
lab"  
/issue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to  
inflorescence development"  
/lab\_host="X10LR"  
/note="Organ: tassels; Vector: HybriZAP; Site:1: EcorI;  
Site:2: XhoI; George Chuck dissected immature tassels  
between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
library in HybriZAP. Sample insert size range was 350 bp  
to 3 kb with a 1 kb average."  
BASE COUNT 155 a 139 c 94 g 131 t  
ORIGIN

Query Match 4.1%; Score 78; DB 136; Length 519;  
Best Local Similarity 99.2%; Pred. No. 2.4e-28;  
Matches 128; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 601 ctgaagtagattcttcattatgatctcgtatcaagaagaagtgagcagtgatg 660  
|||||  
DB 473 ctgaagtagattcttcattatgatctcgtatcaagaagaagtgagcagtgatg 414  
|||||

QY 661 gttatacatcgaaagcaatagtgagcagtgagcagtgagtgatg 720  
|||||  
DB 413 gttatacatcgaaagcaatagtgagcagtgagcagtgagtgatg 354  
|||||

QY 721 gctacaaga 729  
|||||  
DB 353 gctacaaga 345  
|||||

RESULT 12  
BG464635 717 bp mRNA EST 20-MAR-2001  
LOCUS EMI\_72.B07.g1\_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.  
ACCESSION BG464635  
VERSION BG464635.1 GI:13393339  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor

REFERENCE  
AUTHORS Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.  
TITLE An EST database from Sorghum: developing embryos  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions  
below phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: PolyTmX  
High quality sequence start: 23  
High quality sequence stop: 680  
POLYA-No.

FEATURES  
source  
1..717  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Embryo 1 (EMI)"  
/note="Organ: Embryos germinated for 24 hr; Vector:  
pBluescript II from Lambda Zap II; Site:1: XhoI; Site:2:  
EcorI; The library was made from polyA RNA in the cloning  
vector Lambda Zap II. Clones to be sequenced were  
prepared by mass excision."  
BASE COUNT 197 a 127 c 182 g 210 t 1 others  
ORIGIN

Query Match 3.6%; Score 68; DB 154; Length 717;  
Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 958 gaagaattcgaagctgactggtgtgtgttcattatgagtgccacatgtgat 1017  
|||||  
DB 14 gaagaattcgaagctgactggtgtgtgttcattatgagtgccacatgtgat 73  
|||||

QY 1018 atcgatg 1025  
|||||  
DB 74 atcgatg 81  
|||||

RESULT 13  
A1932123 435 bp mRNA EST 02-AUG-1999  
LOCUS 618021C05.x1 618 - Inbred Tassel cDNA Library Zea mays cDNA, mRNA  
DEFINITION sequence.  
ACCESSION A1932123  
VERSION A1932123.1 GI:5670860  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade: Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 435)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu

FEATURES Plate: 618021 row: C column: 05.  
Location/Qualifiers

1. .435

/organism="Zea mays"

/cultivar="Ohio43"

/db\_xref="taxon:4577"

/clone\_lib="618 - Inbred Tassel cDNA Library"

/tissue\_type="tassel"

/dev\_stage="tassel length from 0.1 to 2.5 cm"

/lab\_host="XLOLR"

/note="Organ: tassel; Vector: PAD-GAL4-2.1 (Hybridap); Inbred tassel library from Schmidt lab"

BASE COUNT 109 a 114 c 92 g 120 t

ORIGIN

Query Match 3.4%; Score 65; DB 103; Length 435;  
Best Local Similarity 100.0%; Pred. No. 9.7e-22;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1249 gatgaagaagctgtctattttagcgagatcgacagacacattttagacatgct 1308  
|||||

Db 404 GATGAAGAAGCTGCTATTATTAGCGACATCGACACACATCTTTGAGACATGCT 345

Oy 1309 atcca 1313  
|||||

Db 344 ATTCA 340

RESULT 14

LOCUS BG049561

DEFINITION EM1\_5.G05.g1\_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA

ACCESSION BG049561

VERSION BG049561.1 GI:12501394

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

TITLE 1 (bases 1 to 465)

JOURNAL Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.

COMMENT An EST database from Sorghum: developing embryos

Unpublished (2000)

Contact: Cordonnier-Pratt MM

Department of Botany

The University of Georgia

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions

below phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: T7

High quality sequence start: 73

High quality sequence stop: 450

POLYA=Yes.

Location/Qualifiers

1. .465

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Embryo 1 (EM1)"

/note="Organ: Embryos germinated for 24 hr; Vector: pluscript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from polyA-RNA in the cloning

vector lambda Zap II. Clones to be sequenced were

prepared by mass excision."

BASE COUNT 116 a 96 c 120 g 133 t

ORIGIN

Query Match 3.4%; Score 65; DB 173; Length 465;  
Best Local Similarity 100.0%; Pred. No. 9.7e-22;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1249 gatgaagaagctgtctattttagcgagatcgacagacacattttagacatgct 1308  
|||||

Db 29 GATGAAGAAGCTGCTATTATTAGCGACATCGACACACATCTTTGAGACATGCT 88

Oy 1309 atcca 1313  
|||||

Db 89 ATTCA 93

RESULT 15

LOCUS AM566092

DEFINITION AM566092 503 bp mRNA EST 10-MAR-2000

660061G12.Y1 660 - Mixed stages of anther and pollen Zea mays cDNA,

mRNA sequence.

ACCESSION AM566092 GI:7227451

VERSION AM566092.1

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

TITLE 1 (bases 1 to 503)

JOURNAL Walbot, V.

COMMENT Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Place: 660061 row: G column: 12.

Location/Qualifiers

1. .503

/organism="Zea mays"

/cultivar="Ohio43"

/db\_xref="taxon:4577"

/clone\_lib="660 - Mixed stages of anther and pollen"

/tissue\_type="whole premeiotic anthers to pollen shed"

/dev\_stage="premeiotic anthers to pollen shed"

/lab\_host="XLOLR"

/note="Organ: anthers; Vector: lambda Zap; Site 1: EcoRI;

Site 2: XhoI; Anther and pollen cDNA library.

Directionally sequenced with 5' end at the EcoRI site.

Created by Amie Franklin."

BASE COUNT 146 a 103 c 135 g 119 t

ORIGIN

Query Match 3.4%; Score 65; DB 117; Length 503;  
Best Local Similarity 100.0%; Pred. No. 9.6e-22;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1249 gatgaagaagctgtctattttagcgagatcgacagacacattttagacatgct 1308  
|||||

Db 144 GATGAAGAAGCTGCTATTATTAGCGACATCGACACACATCTTTGAGACATGCT 203

Oy 1309 atcca 1313  
|||||

Db 204 ATTCA 208

Wed Nov 14 08:34:07 2001

us-09-589-510-3.oli.rst

Page 10

1 1 1

0



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2001, 06:55:54 ; Search time 23.2 Seconds  
(without alignments) 1188.962 Million cell updates/sec

Title: US-09-589-510-4

Perfect score: 2263

Sequence: 1 MRFEVSTSKKORIAHTH.....YLDKSSARLQGEQERYIT 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq\_0601:\*

- 1: /SIDSL/gcgdata/geneSeq/geneSeq/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneSeq/geneSeq/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneSeq/geneSeq/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneSeq/geneSeq/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneSeq/geneSeq/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneSeq/geneSeq/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneSeq/geneSeq/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneSeq/geneSeq/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneSeq/geneSeq/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneSeq/geneSeq/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneSeq/geneSeq/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneSeq/geneSeq/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneSeq/geneSeq/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneSeq/geneSeq/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneSeq/geneSeq/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneSeq/geneSeq/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneSeq/geneSeq/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneSeq/geneSeq/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneSeq/geneSeq/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneSeq/geneSeq/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneSeq/geneSeq/AA2000.DAT:\*
- 22: /SIDSL/gcgdata/geneSeq/geneSeq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2263	100.0	455	22	AAV72563
2	2253	99.6	455	22	AAV72562
3	2222	98.2	455	22	AAV72564
4	2211.5	97.7	456	22	AAV72565
5	2157.5	95.3	438	22	AAV72566
6	1994	88.1	458	21	AAAG20673
7	1787	79.0	401	21	AAAG20674
8	1761	77.8	395	21	AAAG20675
9	1723.5	76.2	456	20	AAW74417
10	1722.5	76.1	456	20	AAW74416
11	1721.5	76.1	456	20	AAW93945

12	1541.5	68.1	418	20	AAAV73986
13	973	43.0	463	21	AAAB21093
14	932.5	41.2	471	21	AAAB21094
15	630.5	27.9	334	22	AAAB06111
16	538.5	23.8	242	20	AAAY04324
17	494.5	21.9	128	21	AAAB27851
18	485	21.4	119	21	AAAB27847
19	481	21.3	119	21	AAAB27848
20	477	21.1	121	21	AAAB27849
21	470	20.8	121	21	AAAB27850
22	446	19.7	127	21	AAAB27852
23	434.5	19.2	233	20	AAAB86303
24	364	16.1	161	20	AAAY04309
25	258	11.4	85	21	AAAG02092
26	231	10.2	106	21	AAAG04033
27	135.5	6.0	603	21	AAAG29815
28	135.5	6.0	627	21	AAAG29814
29	132.5	5.9	826	21	AAAY95050
30	127.5	5.6	555	22	AAAB31943
31	126	5.6	2039	19	AAAB56322
32	125	5.5	566	20	AAAY9068
33	125	5.5	566	22	AAAB31932
34	121.5	5.4	391	21	AAAY5245
35	121.5	5.4	391	21	AAAY5246
36	120.5	5.3	391	21	AAAY5244
37	120.5	5.3	644	21	AAAY08320
38	119.5	5.3	444	20	AAAY08327
39	118.5	5.2	918	20	AAAY35654
40	117.5	5.2	440	17	AAAB94600
41	117	5.2	806	21	AAAB05879
42	117	5.2	806	22	AAAB59399
43	115.5	5.1	632	18	AAAB56684
44	115.5	5.1	680	21	AAAG42129
45	115.5	5.1	821	21	AAAG42128

#### ALIGNMENTS

RESULT	1	
ID	AAV72563	standard; Protein: 455 AA.
XX	AAV72563:	
AC	AAV72563:	
XX		
DT	02-MAY-2001	(first entry)
XX		
DE	Maize RuvB orthologue protein #2.	
XX		
KW	Maize; RuvB orthologue; branch migration; heteroduplex extension; homologous recombination; transformatio; transgenic plant.	
XX		
OS	Zea mays.	
XX		
Key		Location/Qualifiers
FT	Region	/note= "High homology region"
FT	Region	214..294
FT	Region	/note= "Putative heptad repeat region"
FT	Region	297..439
FT	Region	/note= "High homology region"
FT	Binding-site	297..305
FT		/label= "Walker Box B
FT		/note= "ATP binding motif"
XX		
XX	W0200105975-A1.	
XX		
XX	25-JAN-2001.	
XX		
XX	13-JUN-2000; 2000WO-US16271.	
XX		
XX	16-JUL-1999; 99US-0144112.	
XX		

Human prostate tum  
Human HEL50 DNA he  
Yeast HEL50 DNA he  
Environmental stre  
Fragment of human  
Sequence homology  
Sequence homology  
Protein fragment e  
Sequence homology  
Protein fragment e  
Protein fragment e  
Secreted protein e  
Human secreted pro  
Human secreted pro  
Arabidopsis thalia  
Arabidopsis thalia  
Candida albicans p  
Amino acid sequenc  
Haemophilus paraga  
Amino acid sequenc  
Amino acid sequenc  
Neisseria meningit  
Neisseria meningit  
Neisseria gonorrhoe  
E. coli FISH prote  
Human CBFA1E10 pro  
Amino acid sequenc  
S4 protein. Homo  
Human translatio  
Protein tyrosine p  
H. pylori ORF 02ge  
Arabidopsis thalia  
Arabidopsis thalia

PA (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PB:

XX WPI: 2001-159537/16.

DR N-PSDB: AAD02567.

PT Novel maize RuVB nucleic acid useful for modulating levels of maize  
RuVB in plants, as probes or amplification primers in the detection,  
quantitation or isolation of gene transcripts -

PS Claim 12: Page 62-63; 87pp: English.

XX The present sequence is a Zea mays RuVB orthologue protein #2. RuVB  
CC along with RuVA catalyses the branch migration process, also known as  
CC heteroduplex extension, in homologous recombination. RuVB is used for  
CC the control of homologous recombination or transformation efficiency in  
CC transgenic plants. The RuVB nucleotide may be used as probes or  
CC amplification primers for detecting, quantifying or isolating gene  
CC transcripts, in detecting deficiencies in the mRNA level during screening  
CC for desired transgenic plants, for detecting gene mutations, for  
CC monitoring upregulation of expression or changes in enzyme activity, for  
CC detecting any number of allelic variants, orthologues or paralogues of  
CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
CC also be used for recombinant expression of its encoded polypeptide, or  
CC for use as immunogen in preparing and/or screening of antibodies, and in  
CC sense or antisense suppression of one or more genes in a host cell,  
CC tissue or plant. The RuVB proteins may be used in assays to agonise or  
CC antagonise the enzyme function, or as immunogens or antigens for  
CC screening antibodies.

SO Sequence 455 AA:

Query Match 100.0%; Score 2263; DB 22; Length 455;

Best Local Similarity 100.0%; Pred. No. 2.8e-190;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIEVOSTSKKORATPHTHIKGLDANGMAIALAGVGQAAREAGLAVDMIRQKK 60  
DB 1 mlreevgstskkgrlaththikgldangmataaafvgqaaareaglavdmirqkk 60  
QY 61 MAGRAVLGAPRATKKTALIGIAOELGSKVPFCPMWGSSEVSSVKKTEVLMEMFRRAI 120  
DB 61 megrevllagppatgkftalaiglaqelsgkvpfcpmvsevssvkktevlmenfrrai 120  
QY 121 GLRIKENKEVGEVTELSPEAEESTGTGYAKSISHVITISLKTGVTKOLKIDSSITDAL 180  
DB 121 glrickenkevgtevtelspeaeestgtgyaksishvitslktvgtkqlkiddsiyal 180  
QY 181 IREKVAAGDVITYEANSAGVKKVGRCDSPAFETYDLAEVYPDPKGEVHKKEIVQDVTL 240  
DB 181 ikekvavagdvityeansagvkkvgrcdspafetydlaeeyvpdpgevhkkkeivqdvtl 240  
QY 241 HLDLDAANQPOGGDILSLMGOMMKPRKTEITKROELNKVYNYIDEGIAELVPGVLF 300  
DB 241 hldldaannqpggqdlslmgommkprkteitkrlgeinlvvnyidegiaelvpvlylf 300  
QY 301 IDEVHMLDIECFSYLNRALESPLSPVILATNRGICNVRGDMTSPHGIPVDDLRLVYI 360  
DB 301 idevhlmdiectsylnrlesplspvillatnrgicnvrgclmvsphgipvdllrlvyl 360  
QY 361 RTEYVGPTEMIQILAIRAOVEIDMDESLAYLGEIGOOTSLURHAIDLISPAVVSKTNG 420  
DB 361 rteyvgptemiqilairaoveidmdeeslaylgeigqtslnraihqispasvvsctng 420  
QY 421 REKICKADLEVSGLYLAKSSARLLQEQERYIT 455  
DB 421 rekickadeevsglylakssarllqeqeryit 455

RESULT 2  
AAV72562

ID AAV72562 standard; Protein: 455 AA.

XX AAV72562;

XX 02-MAY-2001 (first entry)

XX Maize RuVB orthologue protein #1.

KW Maize; RuVB orthologue; branch migration; heteroduplex extension;  
homologous recombination; transformation; transgenic plant.

XX Zea mays.

XX Key Location/Qualifiers

FT Region /note="High homology region"

FT Region 214..294 /note="Putative heptad repeat region"

FT Region 297..439 /note="High homology region"

FT Binding-site 297..305 /label="Walker Box B"

FT /note="ATP binding motif"

XX W0200105975-A1.

XX 25-JAN-2001.

XX 13-JUN-2000; 2000WO-US16271.

XX 16-JUL-1999; 99US-0144112.

XX (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PB;

XX WPI: 2001-159537/16.

XX N-PSDB: AAD02566.

PT Novel maize RuVB nucleic acid useful for modulating levels of maize  
RuVB in plants, as probes or amplification primers in the detection,  
quantitation or isolation of gene transcripts -

PS Claim 12: Page 62-63; 87pp: English.

XX The present sequence is a Zea mays RuVB orthologue protein #1. RuVB  
CC along with RuVA catalyses the branch migration process, also known as  
CC heteroduplex extension, in homologous recombination. RuVB is used for  
CC the control of homologous recombination or transformation efficiency in  
CC transgenic plants. The RuVB nucleotide may be used as probes or  
CC amplification primers for detecting, quantifying or isolating gene  
CC transcripts, in detecting deficiencies in the mRNA level during screening  
CC for desired transgenic plants, for detecting gene mutations, for  
CC monitoring upregulation of expression or changes in enzyme activity, for  
CC detecting any number of allelic variants, orthologues or paralogues of  
CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
CC also be used for recombinant expression of its encoded polypeptide, or  
CC for use as immunogen in preparing and/or screening of antibodies, and in  
CC sense or antisense suppression of one or more genes in a host cell,  
CC tissue or plant. The RuVB proteins may be used in assays to agonise or  
CC antagonise the enzyme function, or as immunogens or antigens for  
CC screening antibodies.

SO Sequence 455 AA:

Query Match 99.6%; Score 2253; DB 22; Length 455;

Best Local Similarity 99.3%; Pred. No. 2.1e-189;

Matches 452; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRIEVOSTSKKORATPHTHIKGLDANGMAIALAGVGQAAREAGLAVDMIRQKK 60  
DB 1 mlreevgstskkgrlaththikgldangmataaafvgqaaareaglavdmirqkk 60

```

QY 61 MAGRAVLGAPATGKTALALGIAOELGSKVPCPMVGSEVSEVKKTEVLMENFRRAI 120
DB 61 magraavllvppatgktalalaglaqelgskvpcpmvgsevsykktevlmenfral 120
QY 121 GLRIKENKEVEGEVTELSPEAEESTTGGYAKSISHVILSLKTVKGTQKLDSSIYDAL 180
DB 121 glrikenkevegevtelspeaeesttggysakshivilslktvkgtkldssiydal 180
QY 181 IKEKVAAGDVIIYIANSAGAVKRVGRCDSPATEVDLEAEVVPPIPKGEVHKKEIVQDVTL 240
DB 181 ikekvaagdviiyiansagavkrvgrcdspatevdlleaevvpipkgevhkkelvqdvtl 240
QY 241 HDLDAANAOPGOGDILSLMGOMMKPRKTEITTEKLRQEIKNKVNRVYIDEGIAELVPGVLF 300
DB 241 hlddaanaopgogdilsimgommkprkteittekrlrqeinkvnrvyidegiaelvpgvlf 300
QY 301 IDEVHMLDIECFSYLNRALSPISPIVILATNRGICNVGRTDMTSPHGIPVDLDRLVII 360
DB 301 idevhmldiecfstylnralsplspivilatnrgicnvgtdmtsphgipvdldrlvii 360
QY 361 RTETVPTMIIQILAIRAOVEIDMDESLAYIGEIGQTSIRHAIQLISPAVSYSKTING 420
DB 361 rtevtptmiiqilairaoveidmdeeslayigeigqtsirhaiqlispasvyskting 420
QY 421 REKICKADLEEVSGLYIDAKSSARLLQEOOERYIT 455
DB 421 rekickadleevsglyidakssarllqeqeryit 455

RESULT 3
AAV72564 standard; Protein; 455 AA.
ID AAV72564
AC AAV72564;
XX 02-MAY-2001 (first entry)
DT
DE Maize RuVB orthologue protein #3.
XX
KW Maize; RuVB orthologue; branch migration; heteroduplex extension;
KW homologous recombination; transformation; transgenic plant.
XX
OS Zea mays.
XX
FH Location/Qualifiers
FT Region 40..86
FT /note="High homology region"
FT Region 214..294
FT /note="Putative heptad repeat region"
FT Region 297..439
FT /note="High homology region"
FT Binding-site 297..305
FT /label= Walker_Box_B
FT /note="ATP binding motif"
XX
PN W0200105975-A1.
XX
XX 25-JAN-2001.
XX
XX 13-JUN-2000; 2000WO-US16271.
XX
XX 16-JUL-1999; 99US-0144112.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPI: 2001-159537/16.
XX
XX N-PSDB; AAD02366.
XX
XX Novel maize RuVB nucleic acid useful for modulating levels of maize
XX RuVB in plants, as probes or amplification primers in the detection,
XX

```

```

PT quantitation or isolation of gene transcripts -
XX Claim 12; Page 62-63; 87pp; English.
PS
XX The present sequence is a Zea mays RuVB orthologue protein #3. RuVB
CC along with RuVA catalyses the branch migration process, also known as
CC heteroduplex extension, in homologous recombination. RuVB is used for
CC the control of homologous recombination or transformation efficiency in
CC transgenic plants. The RuVB nucleotide may be used as probes or
CC amplification primers for detecting, quantifying or isolating gene
CC transcripts, in detecting deficiencies in the mRNA level during screening
CC for desired transgenic plants, for detecting gene mutations, for
CC monitoring upregulation of expression or changes in enzyme activity, for
CC detecting any number of allelic variants, orthologues or paralogues of
CC the gene, or for site directed mutagenesis in eukaryotic cells. It may
CC also be used for recombinant expression of its encoded polypeptide, or
CC for use as immunogen in preparing and/or screening of antibodies, and in
CC sense or antisense suppression of one or more genes in a host cell,
CC tissue or plant. The RuVB proteins may be used in assays to agonise or
CC antagonise the enzyme function, or as immunogens or antigens for
CC screening antibodies.
XX
SQ Sequence 455 AA:
XX
Query Match 98.2%; Score 2222; DB 22; Length 455;
Best Local Similarity 97.6%; Pred. No. 1,le-186;
Matches 444; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 1 MRLEEVOSTSKRKQIATHTHRIKGLGDANGMAIALAGFVGQAARAAAGLAVDMIROKK 60
DB 1 mrleevostskrkqiathtthrikglgdangmalaagfvgqaaraaaglavdmirgk 60
QY 61 MAGRAVLGAPATGKTALALGIAOELGSKVPCPMVGSEVSEVKKTEVLMENFRRAI 120
DB 61 magraavllvppatgktalalaglaqelgskvpcpmvgsevsykktevlmenfral 120
QY 121 GLRIKENKEVEGEVTELSPEAEESTTGGYAKSISHVILSLKTVKGTQKLDSSIYDAL 180
DB 121 glrikenkevegevtelspeaeesttggysakshivilslktvkgtkldssiydal 180
QY 181 IKEKVAAGDVIIYIANSAGAVKRVGRCDSPATEVDLEAEVVPPIPKGEVHKKEIVQDVTL 240
DB 181 ikekvaagdviiyiansagavkrvgrcdspatevdlleaevvpipkgevhkkelvqdvtl 240
QY 241 HDLDAANAOPGOGDILSLMGOMMKPRKTEITTEKLRQEIKNKVNRVYIDEGIAELVPGVLF 300
DB 241 hlddaanaopgogdilsimgommkprkteittekrlrqeinkvnrvyidegiaelvpgvlf 300
QY 301 IDEVHMLDIECFSYLNRALSPISPIVILATNRGICNVGRTDMTSPHGIPVDLDRLVII 360
DB 301 idevhmldiecfstylnralsplspivilatnrgicnvgtdmtsphgipvdldrlvii 360
QY 361 RTETVPTMIIQILAIRAOVEIDMDESLAYIGEIGQTSIRHAIQLISPAVSYSKTING 420
DB 361 rtevtptmiiqilairaoveidmdeeslayigeigqtsirhaiqlispasvyskting 420
QY 421 REKICKADLEEVSGLYIDAKSSARLLQEOOERYIT 455
DB 421 rekickadleevsglyidakssarllqeqeryit 455

RESULT 4
AAV72565 standard; Protein; 456 AA.
ID AAV72565
AC AAV72565;
XX 02-MAY-2001 (first entry)
DT
DE Maize RuVB orthologue protein #4.
XX
XX Maize RuVB orthologue; branch migration; heteroduplex extension;
XX

```

homologous recombination; transformation; transgenic plant.

Zea mays.

Key Location/Qualifiers

Region 41..87

Region /note="High homology region"

Region 215..295

Region /note="Putative heptad repeat region"

Region 298..440

Region /note="High homology region"

Binding-site 298..306

Binding-site /label="Walker\_Box\_B"

Binding-site /note="ATP binding motif"

MO200105975-A1.

25-JAN-2001.

13-JUN-2000; 2000MO-US16271.

16-JUL-1999; 99US-0144112.

(PION-) PIONEER HI-BRED INT INC.

Mahajan PB;

MP1: 2001-159537/16.

N-PSDB; AAD02569.

Novel maize RuVb nucleic acid useful for modulating levels of maize RuVb in plants, as probes or amplification primers in the detection, quantitation or isolation of gene transcripts -

Claim 12: Page 62-63; 87pp: English.

The present sequence is a Zea mays RuVb orthologue protein #4. RuVb along with Ruva catalyses the branch migration process, also known as heteroduplex extension, in homologous recombination. RuVb is used for the control of homologous recombination or transformation efficiency in transgenic plants. The RuVb nucleotide may be used as probes or amplification primers for detecting, quantifying or isolating gene transcripts, in detecting deficiencies in the mRNA level during screening for desired transgenic plants, for detecting gene mutations, for monitoring upregulation of expression or changes in enzyme activity, for detecting any number of allelic variants, orthologues or paralogues of the gene, or for site directed mutagenesis in eukaryotic cells. It may also be used for recombinant expression of its encoded polypeptide, or for use as immunogen in preparing and/or screening of antibodies, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The RuVb proteins may be used in assays to agonise or antagonise the enzyme function, or as immunogens or antigens for screening antibodies.

Sequence 456 AA;

Query Match 97.7%; Score 2211.5; DB 22; Length 456;

Best Local Similarity 97.4%; Pred. No. 9.2e-186;

Matches 444; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

1 MREEVOSTSKKORITHTHIKGLD-ANGMAIALAGVGQAAREAGLAVDMIRK 59

1 mreevostskkqritaththikgldangmampiaagfvyqaareaglavdmirgk 60

60 KMAGRAVLLAGPATGKTALALGIAOELGSKVPFCPMVGESEVYSSEVKTEVLMENFRRA 119

61 kmaagrallagpatgktaalalgiaelgskvpcpmvgeevyssevkktelvmenfrira 120

120 IGLIRKENEVYEGEYTESPEAEESTTGAKKSHVLSIKTYKQKQLKLDSSITDA 179

121 iglirkenevyegeytelspeaeesttgakksishvlsiktykqkqlkldpsiyda 180

180 LIKEKAVAGDYIYIANSAGAVKRVGRCDSPATEYDLAEAEVVPJPKGEVHKKEIVODVT 239

181 likekavagdyilyiensagavkvrgrcdsfateydlaeaevvpjpkgevhkkelvqdv 240

240 LHMDDANAQPOGGODIISLMGOMAKPKRTETKRLQEIINKVYVNRITDEGIAELVPGVL 299

241 lhmddanaqpoggodislmgomakprkrtetkrlqeiinkvynritydegiaelvpgvl 300

300 FLDEVMHMDICFSLYNALLESPLSPYIATNNGICNVGRCDMTSPHGIPVDDLRLVI 359

301 fldevmhdicfslynallesplspyiattngicnvgrcdmtsphgipvdlrlvi 360

360 IRTETVPTMIGIILATRAQVEEIDMDESLAVLGEIGQOTSLRHAQIQLISPASVSKTN 419

361 irtetvptmigiilatraqveeidmdeslavlgeigqotslrhaqiolispasvsktn 420

420 GREKICRADLEEVSGLYIDAKSSARLLQEOERIT 455

421 grekicradleevsglyidakssarllqeoeryit 456

RESULT 5

AAV72566

ID AAV72566 standard; Protein: 438 AA.

AAV72566;

02-MAY-2001 (first entry)

Maize RuVb orthologue protein #5.

Maize: RuVb orthologue; branch migration; heteroduplex extension; homologous recombination; transformation; transgenic plant.

Zea mays.

Key Location/Qualifiers

Region 40..69

Region /note="High homology region"

Region 197..277

Region /note="Putative heptad repeat region"

Region 280..422

Region /note="High homology region"

Binding-site 280..288

Binding-site /label="Walker\_Box\_B"

Binding-site /note="ATP binding motif"

MO200105975-A1.

25-JAN-2001.

16-JUL-1999; 99US-0144112.

13-JUN-2000; 2000MO-US16271.

MP1: 2001-159537/16.

N-PSDB; AAD02570.

Novel maize RuVb nucleic acid useful for modulating levels of maize RuVb in plants, as probes or amplification primers in the detection, quantitation or isolation of gene transcripts -

Claim 12: Page 62-63; 87pp: English.

The present sequence is a Zea mays RuVb orthologue protein #5. RuVb along with Ruva catalyses the branch migration process, also known as heteroduplex extension, in homologous recombination. RuVb is used for the control of homologous recombination or transformation efficiency in transgenic plants. The RuVb nucleotide may be used as probes or

CC amplification primers for detecting, quantifying or isolating gene  
CC transcripts, in detecting deficiencies in the mRNA level during screening  
CC for desired transgenic plants, for detecting gene mutations, for  
CC monitoring upregulation of expression or changes in enzyme activity, for  
CC detecting any number of allelic variants, orthologues or paralogues of  
CC the gene, or for site directed mutagenesis in eukaryotic cells. It may  
CC also be used for recombinant expression of its encoded polypeptide, or  
CC for use as immunogen in preparing and/or screening of antibodies, and in  
CC sense or antisense suppression of one or more genes in a host cell,  
CC tissue or plant. The Ruvb proteins may be used in assays to agonise or  
CC antagonise the enzyme function, or as immunogens or antigens for  
CC screening antibodies.

XX Sequence 438 AA:

Query Match 95.3%; Score 2157.5; DB 22; Length 438;  
Best Local Similarity 96.3%; Pred. No. 4.8e-181;  
Matches 438; Conservative 0; Mismatches 0; Indels 17; Gaps 2;

QY 1 MREEVOSTSKORIRATHTHIKGLDANGMATAALAAEFGQAAREAGLAVDMTRKK 60  
DB 1 mreevgstskqrllaththikgldangmalalaagfvq-----gkk 44  
QY 61 MAGRAVLLAGPPATGKTALALIGTQELGSKVPPCPMGVSEVSEVKKTEVIMENFRRAI 120  
DB 45 magravllagppatgktala-glaqelgskvppcpmgvsevsevkktevimenfral 103  
QY 121 GRIKENKEVEGETELSPPEASTTGGYAKSISHVILSKTVKGTQLDLSIYDAL 180  
DB 104 grikenevegevetelspeeasttgggyakshvilsktkvgtqkldlsiydal 163  
QY 181 IKEKVAVDVITYEANSQAVRVGRCDSFATERYDLEAEVYPIPKGEVHKKEIVODVTL 240  
DB 164 ikekvavdvityeansqavrvgrcdsfaterydleaevyppikgevhkkeitvovtl 223  
QY 241 HDLDANNAPOGGDILSLMGOMMKPRKTEETKLRQENKVVARYIDEIGIAELVPGVLF 300  
DB 224 hlddanagppggdilsimgommkprkteetkrlrqenkvvarryidegiaelvpgvlf 283  
QY 301 IDEVHMLDIECFSYLNRALESPLSPVILATNRGICNVKGTDMTSPHGIFVLDLRLVIT 360  
DB 284 idevhlmdiectsylnralesplspvillatnrgicnvrgtdmtspbgifvldlrlvll 343  
QY 361 RRETYGPTETMIQILAIRAQVEEDIMDESLAYLGEIGQOTSRLRAIOLISPAVSKTNG 420  
DB 344 rretygptetmiqlairaqveeideeslaylgeigqotsrlraiolispavsktng 403  
QY 421 RERICKADLEEVSGLYLDAKSSARLLOOERYTT 455  
DB 404 rekickadleevsgilyldakssarllloeqerytt 438

RESULT 6

AAG20673 standard: Protein, 438 AA.

AAG20673;

17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 22955.

KW Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

XX EPI033405-r2.

XX 06-SEP-2000.

XX

PF 25-FEB-2000: 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 26-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 14-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134941.  
PR 19-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0135259.  
PR 21-MAY-1999; 99US-0135629.  
PR 24-MAY-1999; 99US-0136021.  
PR 25-MAY-1999; 99US-0136382.  
PR 27-MAY-1999; 99US-0136782.  
PR 28-MAY-1999; 99US-0137222.  
PR 01-JUN-1999; 99US-0137528.  
PR 03-JUN-1999; 99US-0137502.  
PR 04-JUN-1999; 99US-0137724.  
PR 07-JUN-1999; 99US-0138094.  
PR 08-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 10-JUN-1999; 99US-0139453.  
PR 14-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139454.  
PR 17-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 22-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.

PR	28-SEP-1999;	9905-0156458.
PR	29-SEP-1999;	9905-0156596.
PR	04-OCT-1999;	9905-0157117.
PR	05-OCT-1999;	9905-0157753.
PR	06-OCT-1999;	9905-0157865.
PR	07-OCT-1999;	9905-0158029.
PR	08-OCT-1999;	9905-0158232.
PR	12-OCT-1999;	9905-0158369.
PR	13-OCT-1999;	9905-0159293.
PR	13-OCT-1999;	9905-0159294.
PR	13-OCT-1999;	9905-0159295.
PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	25-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

Query Match	88.1%;	Score 1994;	DB 21;	Length 458;
Best Local Similarity	85.1%;	Pred. No. 1.1e-166;		
Matches 387;	Conservative 39;	Mismatches 29;	Indels 0;	Gaps 0

QY	1	KRIEVOSTSKORIANHHIKSLIGDANGMANIALAAGFVQGAARAGLAVDMTRCK	60
Db	4	vkleeigstakqriahtnchikjigleptgkrliaagfyqgleateaaqlvwdmllkqk	63
QY	61	MAGRAVLLAGPATGKTALALJAOLGSKVPCPMGSEVSEYSEVKTGLVLENFRRI	12
Db	64	magkalllaagppgkrtalajgsqelgskvpcpmgseysevktlevmentrral	12
QY	121	GLRIKEMKYEEBEVVELSPREAEESTGGYAKSISHVILSKTYVGTOLKLDSTYDAL	18
Db	124	gltlkekeveyeeveelspreetestlgygksjshvltlktvkgtkhklkldptlydal	18
QY	181	IKERVAADVITYLEANSQAVKRYGDCSPAEYDLAEAEVPIPKCEVHKKEIVODWTL	24
Db	184	lkekvavgdvityleansgavkrygsdafeatefdleeeeyvpipkcevhkkkeivodvtl	24
QY	241	HDLDAAAOQGGODILSLMGOMKPRKTELTTELKROELIKVNVYRTIDGIELVPGVLF	30
Db	244	qdlldaaanarqggqdlslsmgumkprkteletldlreghkvnvryldagvaelvpgvlf	30
QY	301	IDEVHMLDIECFEYLRALAESPLSYVLATNRRICVNRGDTQSPHGIPVOLLRLVIT	36
Db	304	ldeyhmldmecefsylrnlrleslsplvlfatmgvclvrgtdmpshyprldlrlviti	36
QY	361	RTEYFGPTEKIQLLAIRAOVEIDMDESLAYGELTQQTSLRHAQTILSPASVSKTNG	42
Db	364	rtqyfpsemiqilairagveelivdeecvlylgelqgrtslrthavqllspasvsktng	42
QY	421	REKIKACKADLEEVSGYLDAKSSARILQEOQERYIT	45
Db	424	rdnkckadleevtsylldakssakllneqekyis	45

RESULT 7  
AAG20674  
ID AAG20674 standard; Protein; 401 AA.  
XX  
AC AAG20674;  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22956.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155658.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 79.0%; Score 1787; DB 21; Length 401;

Best Local Similarity 86.3%; Pred. No. 1.4e-148;

Matches 346; Conservative 33; Mismatches 22; Indels 0; Gaps 0;

QY 55 MIROKKNAGRAVLLAGPATGKTALALGIAOELGSKVPCFPMWSEVYSSEVKKTEVIME 114  
DB 1 mikykkmagkalllaagpgygtkcalalgtlsgelgskvpcfgmwygevsyevkktevlme 60  
QY 115 NFRRAIGLRIRKEKEVEGEVTELSPEARSTTGCTAKSISHVITSLKTYVKGTKQLKLD 174  
DB 61 nfrtraiglrirkekevegevtelspeetstltgygyskshvltlktvkgtkhklldp 120  
QY 175 SIYDALIKKEVAVAGDVYIIFANGSAVKRVCDFATEVDLEEEVVPPIPKGVHKKKEI 234  
DB 121 tlydallikevavagdvlyiifangsaavkryvrsdaftateidleeeyvpiipkgevnkkkei 180

QY 225 VODVTLHDLDAANAPOGGDIILSLMGOMKPRKTEITKLEROINKVNRXIDEIAEL 294  
DB 181 vgdvtrlqlddaanaarp9ggdilislmqgmkpkrkteitldkrlrgelnvnrxyidegvael 240  
QY 295 VPGVLEIDEVHMDIECFESYLNRALESPLSVTLATNRGICNRCGDMKSPHGIPD 354  
DB 241 vpgvlfidevhlmdiescfesynralesslpsvltatnrgicnrcgdmkspgipd 300  
QY 355 DRVYIRRTETGPTPEMQLIARQVEIEMDESLAYLGEIGQOTSLEKRAIQLISPA 414  
DB 301 drvylirtqlydpsemqqlalraqvleltvdeclvlgelqgrtslrnavqlispasl 360  
QY 415 VSKTNGREKICKADLEEVSGLYIDAKSSARLLPOOGRYIT 455  
DB 361 vakmgtrdnickadieevtslyidakssakllneqgekyls 401  
  
RESULT 8  
AAG20675  
ID AAG20675 standard; Protein; 395 AA.  
XX  
AC AAG20675;  
DT  
DE 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22957.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
PN EP1033405-A2.  
PD  
PD 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134229.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.



PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142970.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148341.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0151930.  
PR 10-SEP-1999; 99US-0152363.  
PR 13-SEP-1999; 99US-0153070.  
PR 15-SEP-1999; 99US-0153758.  
PR 16-SEP-1999; 99US-0154018.  
PR 20-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157173.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161362.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

Query Match	77.88;	Score 1761;	DB 21;	Length 395;
-------------	--------	-------------	--------	-------------

Best Local Similarity 86.3%; Pred. No. 2.6e-146;

Matches	341;	Conservative	32;	Mismatches	22;	Indels	0;	Gaps	0;
---------	------	--------------	-----	------------	-----	--------	----	------	----

Qy	61	MAARAVILLAGPANGKATALAGINOELGSKVPECPMWSGEVSEYKKTUWLMENPRAL	120
		:	
Db	1	magkallilaappgqkatalagisqelgskvpfcipmwygevsyevkveymenfiral	60
Qy	121	GLRIKENKEVEGETELSPPEASTGTGYAKSISHVILSLKVTWKTKQLKLDSDSYDAL	180
		:	
Db	61	griketkeveyegevtelspsetcsileggygkskshviltlckvkgkhhklidplydal	120
Qy	181	IKERAVAVGDVILYIEANGSAVKRVGRCDSPATEYDLAEAEVYLPJKEGVHKKKEIVODVTL	240
		:	
Db	121	ikekavaydvilyieangsavkrvgrsdafatefdeleaeeyvplpkgevhkkkeltvgvtl	180
Qy	241	HDLDAANAPGCGODIISLMGOMMKPKRTETELRQEIINKVYNRYIDDEISIAELVEGVLF	300
		:	
Db	181	qdlidaanaarp9g9gdllslmgqmkkpkrteltldlrlqehlnkvnrlyidegvaaelpvplf	240
Qy	301	IDEVHMLDIECFSVLNLRLAESPLPIYILATNRGICVNRGTDMTSPHGIVDLIDLRLVII	360
		:	
Db	241	idevhlmldecfsvlnrlaeslspylilfatnryvcnrvrgtdmbspbyrgldldlrlvii	300
Qy	361	RTEYVGPTEMTQIILAIRAQVEIDMDESLAVIGETIGQOTSRLRAHLOISPASVASKTNG	420
		:	
Db	301	rtqdydspemigqiairaqveeltvdeecclvyligeltqgrtscrlnavglispasivakmg	360
Qy	421	REKICKADLEEVSGIYLDAKSASALLDQOQRVYT	455
		:	
Db	361	rdnckckadieevtsylldaksasalllneqgkxlys	395

RESULT	9
AAW74417	ID
AAW74417	standard; Protein; 456 AA.
AAW74417	AC
10-MAY-1999	(first entry)
TIP49	protein sequence #2.
TIP49;	TATA binding protein; TBP binding protein;
DNA	transcription control.
Rattus	sp.
MO9900419-A1.	
07-JAN-1999.	
25-JUN-1998;	98WO-JP02836.
27-JUN-1997;	97JP-0187398.
(SUME )	SUMITOMO ELECTRIC IND CO.
Kishimoto T,	Makino Y, Niwa S, Tamura T;
WPI;	1999-095682/08.
N-PSDB;	AAI18199.
TBP-binding	protein with DNA helicase and ATPase activities - gene
encoding	it, and antibodies recognising it.
Claim 2;	Page 41-43; 64pp; Japanese.

CC transcription control by TBP in vivo, and in detection of the blocking of  
CC DNA transcription.

Sequence: 456 AA;

Query Match	Score	DB	Length
76.28;	1723.5;	20;	456;

Best Local Similarity 72.58; Pred. No. 6.3e-143;

Matches	330;	Conservative	73;	Mismatches	51;	Indels	1;	Gaps	1.
---------	------	--------------	-----	------------	-----	--------	----	------	----

Qy	1	MREVEOSTSKXOJIAHHTHKIGJLGDANGALAFVGOAAAREPAALADMIROK	60
Db	1	mkteevstctqrlashtsvktygldeogstakaaaglyvgenataeagyivellskk	60
Qy	61	MAGRAVLLAGPPAPGKTALALGIAOELGSKVPECPMWSEVSEVKKTEVIMENFRAT	120
Db	61	magravlllagppgsktalatalaiaqelsgskvfcfcmwyseystelktevimenfrral	120
Qy	121	GLRIKENKEVEGEVTELSPEEASSTTGGYAKSISHYIISLKYKKCKOKLIDSSITDAL	180
Db	121	gllrketkevegeveteltpceetnmpngy9akltshv1qktakgkqklpdsstiesl	180
Qy	181	IKKEAAVDVITYIEANSQAVKRVKRCSEAFSEVYDLAEAEVPIPKGEVHKKEIYODVTL	240
Db	181	qkerveagdvityieansgavrvtrgrcdtyatefldeeevyprlpkgdvnhkkelqdvtl	240
Qy	241	HDLDAAANQPOGGODIUSLGMOMKPKRTEITEKLRQELNKVNRRTYDEGIAELVPSVLF	300
Db	241	hdlldvaanarpgggd1lsmmgqlmkpkkteltcdk1lrgelnkvnmky1ddgjaaelvpgvlf	300
Qy	301	IDEVHMDIDEEFSLYNRLAESPLPIYALANRRCICVNRGT-DWTSPHGIPVDLIDLVI	359
Db	301	iddevhmdideefcylphralesispiylvfaasnrgncvllrgtedtlspghjpldlldtvm1	360
Qy	360	IRTEYGPTEMIQILAIRAOVEEIDMDSESLAYIGETGOOTSRLHAIQILISPSASVSKTN	419
Db	361	irtmlypqpemkqiklitrageleginiseealnhygelgkrtklrysqvlllpanllakln	420
Qy	420	GREKCRKADLEEVSGLYLDKASRSARLLOEOERYT	454
Db	421	gkdstlekehveeiseltydasssklllaadqdkym	455

```

RESULT 10
AAW74416
ID AAW74416 standard; Protein; 456 AA.
XX
AC AAW74416;
XX
DT 10-MAY-1999 (first entry)
XX
DE TIP49 protein sequence.
XX
KW TIP49; TATA binding protein; TBP binding protein;
KW DNA transcription control.
XX
OS Rattus sp.
XX
PN WO9900419-A1.
XX
PD 07-JAN-1999.
XX
PF 25-JUN-1998; 98WO-JF02836.
XX
PR 27-JUN-1997; 97JP-0187398.
XX
PA (SUME ) SUMITOMO ELECTRIC IND CO.
XX
PI Kishimoto T, Makino Y, Niwa S, Tamura T,
DR WPI; 1999-095682/08.
DR N-PSDB; AAX18198.
XX

```

PT TBP-binding protein with DNA helicase and ATPase activities - gene  
PT encoding it, and antibodies recognising it.

PS Claim 1; Page 37-39; 64pp; Japanese.

XX This sequence is the TATA-binding protein (TBP) binding protein,  
CC designated TBP49, of the invention. TBP49 and its associated nucleic  
CC acids and antibodies are useful in investigation of the process of DNA  
CC transcription control by TBP in vivo, and in detection of the blocking of  
CC DNA transcription.

SO Sequence 456 AA;

Query Match 76.1%; Score 1722.5; DB 20; Length 456;  
Best Local Similarity 72.3%; Pred. No. 7.8e-143;  
Matches 329; Conservative 74; Mismatches 51; Indels 1; Gaps 1;

OY 1 MRIEVOSTSKKQRIATHTHIKGLDANGMALAAGFVQAAREAGLAVDMTRKK 60  
DB 1 mkleevkstktqriashshvkgldesglakgaasglvgenareacgivelleskk 60  
OY 61 MAGRAVLLAGPPATGKTALALGIAOELGSKVPCPMVGSSEYSEVKKTEVLMENFRRAI 120  
DB 61 magravllagppatgktalalalagelsgskvpcpmvgsseystelktevlmennfral 120  
OY 121 GLRIKENKVEYEGEVTELSPEAEESTGTGVAKSISHVIISLKTWKGTKOLKSSITYDAL 180  
DB 121 glriketkeveyegevteltpcelempmgysktishvliigltakgtkqlkldpsifesi 180  
OY 181 IKERKAVGDVYIEFANSKAVKRVGKDSFATEYDLAEAEVYPIPKGEVHKKEIVODVTL 240  
DB 181 qkerveagdvyleansgavkrgvcdyatefdleaeeyvlpdpvdkhkkeliqdvtl 240  
OY 241 HDLDANAPQGGODILSLMGQMKPRKTEITTEKLRQEIKNVKNRYIDEGIAELVPGVLF 300  
DB 241 hldvanarpggqdlismmgqmkprkteitdklrgeinkvnkyldgvaaelvpgvlf 300  
OY 301 IDEVHMLDIECFSTYNRALESPLSPYIVLANNRGICNVKGT-DMTSPHGIPTVDLRLVYI 359  
DB 301 vdevhmldiecfylhraleesslapiyifasnrgncvtrgdedltsphgipdlidrlvml 360  
OY 360 IRTEYGPTEMIQILAIRAAGEIDMEESLAYIGEIGQOTSLRHAQILISPAVSVKTN 419  
DB 361 irtmlypgegmqikliraqteginiseeanlnhgeigtktltlrysvqltpnallakin 420  
OY 420 GREKICKADLEEVSGLYLDANSSARLLQEOQERYI 454  
DB 421 gkdsiekehveiselgydaksakilladgdkym 455

RESULT 11  
ID AAW93945  
AAW93945 standard; Protein; 456 AA.

XX AAW93945;

DT 30-JUN-1999 (first entry)

DE Human regulatory molecule HRM-1 protein.

KW Human regulatory molecule: HRM-1; cytosolic activity; immune modulator;  
KW transcription factor: enhancer; cell proliferation stimulation; cancer;  
KW treatment; microarray; detection; diagnosis; cell proliferation disease;  
KW leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone;  
KW brain; lung; liver; ovary; skin; teratocarcinoma; immune response;  
KW allergy; asthma; diabetes; multiple sclerosis; Grave's disease;  
KW myasthenia gravis.

XX Homo sapiens.  
XX OS  
XX WO9915658-A2.  
XX PN

PD 01-APR-1999.

XX 22-SEP-1998; 98WO-US19839.

XX 23-SEP-1997; 97US-0933750.

XX (INCY-) INCYTE PHARM INC.

PI Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lal P;

PI Shah P, Yue H;

DR WPI: 1999-254710/21.

DR N-PSDB; AAX24059.

XX New human regulatory molecules

XX Claim 1; Page 60-61; 76pp; English.

CC This invention describes novel human regulatory molecules (HRM) which  
CC have cytosolic activity and act as immune modulators, transcription  
CC factors or enhancers. The HRMs can be used to stimulate cell  
CC proliferation. Antagonists and agonists of the proteins of the invention  
CC can be used to treat cancer. The encoding nucleic acids can be used in  
CC microarrays to detect polynucleotides (and their expression levels) that  
CC encode HRMs in a biological sample. The HRMs and microarrays can be used  
CC to diagnose, treat or prevent cell proliferation diseases especially cancer,  
CC e.g. leukemia, lymphoma, myeloma, adenocarcinoma, sarcoma, cancer of e.g.  
CC bladder, bone, brain, lung, liver, ovary, skin, etc. teratocarcinoma, or  
CC to treat or prevent immune responses e.g. allergies, asthma, diabetes,  
CC multiple sclerosis, Grave's disease or myasthenia gravis.

SO Sequence 456 AA;

Query Match 76.1%; Score 1721.5; DB 20; Length 456;  
Best Local Similarity 72.5%; Pred. No. 9.5e-143;  
Matches 330; Conservative 72; Mismatches 52; Indels 1; Gaps 1;

OY 1 MRIEVOSTSKKQRIATHTHIKGLDANGMALAAGFVQAAREAGLAVDMTRKK 60  
DB 1 mkleevkstktqriashshvkgldesglakgaasglvgenareacgivelleskk 60  
OY 61 MAGRAVLLAGPPATGKTALALGIAOELGSKVPCPMVGSSEYSEVKKTEVLMENFRRAI 120  
DB 61 magravllagppatgktalalalagelsgskvpcpmvgsseystelktevlmennfral 120  
OY 121 GLRIKENKVEYEGEVTELSPEAEESTGTGVAKSISHVIISLKTWKGTKOLKSSITYDAL 180  
DB 121 glriketkeveyegevteltpcelempmgysktishvliigltakgtkqlkldpsifesi 180  
OY 181 IKERKAVGDVYIEFANSKAVKRVGKDSFATEYDLAEAEVYPIPKGEVHKKEIVODVTL 240  
DB 181 qkerveagdvyleansgavkrgvcdyatefdleaeeyvlpdpvdkhkkeliqdvtl 240  
OY 241 HDLDANAPQGGODILSLMGQMKPRKTEITTEKLRQEIKNVKNRYIDEGIAELVPGVLF 300  
DB 241 hldvanarpggqdlismmgqmkprkteitdklrgeinkvnkyldgvaaelvpgvlf 300  
OY 301 IDEVHMLDIECFSTYNRALESPLSPYIVLANNRGICNVKGT-DMTSPHGIPTVDLRLVYI 359  
DB 301 vdevhmldiecfylhraleesslapiyifasnrgncvtrgdedltsphgipdlidrlvml 360  
OY 360 IRTEYGPTEMIQILAIRAAGEIDMEESLAYIGEIGQOTSLRHAQILISPAVSVKTN 419  
DB 361 irtmlypgegmqikliraqteginiseeanlnhgeigtktltlrysvqltpnallakin 420  
OY 420 GREKICKADLEEVSGLYLDANSSARLLQEOQERYI 454  
DB 421 gkdsiekehveiselgydaksakilladgdkym 455

RESULT 12  
AAV73986

ID	AAV73986	standard; Protein; 418 AA.
XX		
AC	AAV73986;	
XX		
DT	14-MAR-2000	(first entry)
XX		
DE	Human prostate tumor EST fragment derived protein #173.	
XX		
KW	Pancreas; tumor; EST; expressed sequence tag; human; cytostatic; treatment.	
XX		
OS	Homo sapiens.	
XX		
PN	DE19820190-A1.	
XX		
PD	04-NOV-1999.	
XX		
PF	28-APR-1998; 98DE-1020190.	
XX		
PR	28-APR-1998; 98DE-1020190.	
XX		
PA	(META-) METAGEN GES GENOMEFORSCHUNG MBH.	
XX		
PI	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;	
XX		
DR	WPI; 1999-621386/54.	
XX		
DR	N-PSDB; AA52915.	
XX		
PT	New human nucleic acid sequences from pancreatic tumors, and related proteins -	
XX		
PS	Claim 23; Page 381; 502pp; German.	
XX		
CC	This invention describes novel polypeptides and their encoding nucleic acids derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. AAV7814-Y74252	
CC	represent protein fragments encoded by the human pancreatic tumor CDNA library derived expressed sequence tag (EST) sequences represented in AA52858-Z53014.	
CC		
XX		
SQ	Sequence 418 AA;	
Query Match	68.1%; Score 1541.5; DB: 20; Length 418;	
Best Local Similarity	73.8%; Pred. No. 5.3e-127;	
Matches 296; Conservative	61; Mismatches 43; Indels 1; Gaps 1;	
47	EAAGLAVNDIRQKKMAGRAVLLAGPATKTPALAGIAOELGSKVPECPMGSEYSSRY	106
12	gcyvavlelkkkkmagavllaagpbgtktalalageljskvpfcgmvsseystel	71
107	KTEVLMNEFRRAIGLRIRKENKVEYEGVELSPBEAESTGGVAKSISHVIISLTKYK	166
72	kteylmenfrraiglrirketkeveyegveltptceetnmggykktishvliglktay	131
167	TKOLKIDSSITVALIKKEKAVAGDVITYIENSGAVNRVCGDSFAEYDLEAEYVPIKRG	226
132	tkqlklpsifelslqkerveagdvilyleansgavnrqgrcdtyatefdlaeeypvlpyg	191
227	EYHKKKEIVQDVTLLHDLDAANAPOGGODITSLMGOMMKPRKTEETKLRGINKVVRK	286
192	dvhkkkelllgdvtllhldlanaarpagggdllsmmgqlmkpkteitdklrgelnkvnyk	251
287	IDEGIAELVPGGLFIDEVHMLDIEGFSLNLRALBESPLSYILATNRRGICANRGK-DMTS	345
252	ldggaiaelvpgvlltydevymaldieftylhralestlapvlfasrnqvcvllrgtedlts	311
346	PHGIVNDLRLVLIIRTEYGTETEMIQILAIRAQVEEIDMDPEESLAVYCGEIGOOTSLRHA	405
312	phgavpdlldlrvmlitrlmlyprgemkqikliragteginisesealnghelgltkttlrys	371
406	IQILISPAVSVKTKNGREKICADLDEEVSGLYIDAKSSARLL 446	

[illegible]

```

Db      184 kdkvgagdvitdkatgyskigrsftrardgdamsgtkftvgcpdgelqkrkevhtvs 243
OY      240 LHDLDANAAGOGGDDIISLMGOMKPKRTITEKLEKROEINKVYNRRIIDEGIAELVGVCL 299
Db      244 lhelvlnstrfbg---flal-----fsgdtgkisevreqhlnakvaeewreegkaelipgvl 296
OY      300 FIDEVHMDICEFSYLNALESPPIVILATNRGICNVRGTDMTSPHGIPVDLDRLYI 359
Db      297 fidevhamdiesfslmratesdmaylmatnqrltrltgtsyvspghlpidldlrl 356
OY      360 IRTETVGPTEMIQILAIRAQVEEIDMDEESLAYLGEIGQOTSLRHAIQILISPAVSVEKTN 419
Db      357 vsttprsekdtkqlrlrceeedvemsedaytvltrigletslryaqlilaaalvcrkr 416
OY      420 GREKICRADLEFVSGLYIDAKSSARLLQEOOERYI 454
Db      417 kgtsevqvddikrvtyslfldearstqymkayqdafl 451

```

## RESULT 14

AAB21094  
ID AAB21094 standard; Protein; 471 AA.

```

XX      AC      AAB21094;
XX      DT      22-DEC-2000 (first entry)
XX      DE      Yeast HEL50 DNA helicase.
XX      KW      Yeast; HEL50; DNA helicase; TIP49 homologue; antibody; antisense therapy;
XX      KW      diagnosis; helicase-associated disease; genetic disorder; tumour; cancer.
XX      OS      Saccharomycetes cerevisiae.
XX      PN      WO200047731-A1.
XX      PD      17-AUG-2000.
XX      PF      22-NOV-1999; 99MO-JP06519.
XX      PR      10-FEB-1999; 99JP-0033062.
XX      PA      (SUME ) SUMITOMO ELECTRIC IND CO.
XX      PI      Tamura T;
XX      DR      WPI; 2000-543587/49.
XX      DR      N-PSDB; AAA90317.
XX      PT      DNA helicase of human origin, antibodies to it and DNA encoding it for
XX      PT      investigation diagnosis and treatment of genetic disorders involving
XX      PT      helicase gene expression -
XX      PS      Claim 4; Page 44-45; 49pp; Japanese.
XX      CC      This sequence represents a novel yeast DNA helicase, HEL50, which
XX      CC      has homology with a part of human TIP49. The invention relates to
XX      CC      human and yeast HEL50 DNA helicases (AAB21093, AAB21094), and to cDNAs
XX      CC      encoding them (AAA90316, AAA90317). It also encompasses HEL50 antisense
XX      CC      oligonucleotides and anti-HEL50 antibodies. HEL50 nucleotides, proteins
XX      CC      and antibodies may be used for the investigation, diagnosis and
XX      CC      treatment of diseases with which the DNA helicase is associated, such as
XX      CC      genetic disorders involving variant forms of the gene, and tumours in
XX      CC      which altered levels of expression of the DNA helicase gene occur.
XX      PS      Sequence 471 AA;

```

Query Match 41.2%; Score 932.5; DB 21; Length 471;  
Best Local Similarity 44.6%; Pred. No. 1.8e-73;  
Matches 198; Conservative 88; Mismatches 139; Indels 19; Gaps 6;

```

OY      15 IATHTHKGIGLDANGMATAIAGFVGQAAREAGLAVDMIROKMGAGRAVLIACPPAT 74
Db      20 laashshltglgldentlqprptsegmvgqlqarraagvllkmvqngtlaqraavlvaqppst 79
OY      75 GKTAALAGIAEOEGSVKPPCWPMWSEVSEVSEKTEWLENFPRRAIGLRIKEKEVEEGE 134
Db      80 gktalamvysqslgkdvplfaigseifslsktealqaftrksigikikeetellege 139
OY      135 VTELSPDEASTTGGYAKASISHVIISLKTGVKTRQKLDSSIYDALIKERKAVGVYIYE 194
Db      140 vveiqid--rstlqghkqg---klfiktmetiylgnkmi dglttkekylagdvtsid 193
OY      195 ANGAARKRGRCSFATVDLEA-----EEVPIPKGEVKKKKEIVODVTLLHDLDAANAQP 250
Db      194 kasgkltiklgr--stasrtdydamgadrltfvgcpegelqkrktvhtvs lhelvlnstr 251
OY      251 OGGODILSLMGOMKPKRTITEKLEKROEINKVYNRRIIDEGIAELVGVFIDEVHMDIOTE 310
Db      252 qg---flal-----ftgdqgelrsevrtdqlnkvaewkeegkaeivpgvl fidevmdlde 304
OY      311 CFSYLNRALESPLSPIVILATNRGICNVRGTDMTSPHGIPVDLDRLYIIRTEVGPTEEM 370
Db      305 cfsflmratedefaplvmmatnrgvsktrgtnysphgipdlldrsllttksynegei 364
OY      371 IQILAIRAQVEEIDMDEESLAYLGEIGQOTSLRHAIQILISPAVSVEKTN 430
Db      365 ktlslraqegeevelssdailltkrgetslrlysnllsvaqilamkrkntveevdk 424
OY      431 EVSGLYIDAKSSARLLQEOOERYI 454
Db      425 rayllflldsarsvkynenesqyl 448

```

## RESULT 15

AAB80611  
ID AAB80611 standard; Protein; 334 AA.

```

XX      AC      AAB80611;
XX      DT      02-MAY-2001 (first entry)
XX      DE      Environmental stress tolerant protein SEQ ID 8.
XX      KW      Environmental stress resistance; salt; heat; desert; transgenic plant.
XX      KW      Bruguiiera sezangula.
XX      OS      Bruguiera sezangula.
XX      PN      WO200106006-A1.
XX      PD      25-JAN-2001.
XX      PF      19-JUL-2000; 2000MO-JP04862.
XX      PR      19-JUL-1999; 99JP-0235910.
XX      PR      24-MAR-2000; 2000JP-0085377.
XX      PA      (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX      PI      Yamada A, Ozeki Y, Saito T;
XX      DR      WPI; 2001-147355/15.
XX      DR      N-PSDB; AAF74190.
XX      PT      Screening method to obtain DNA encoding environmental stress resistance
XX      PT      factor, useful for producing transgenic plants resistant to
XX      PT      environmental stress -
XX      PS      Claim 22; Page 86-87; 167pp; Japanese.
XX      CC      Polynucleotide sequences AAF74187 - AAF74218 encode proteins
XX      CC      AAB80608 - AAB80639, which impart environmental stress resistance. The
XX      CC      invention relates to a method for identifying DNA encoding proteins
XX      CC      imparting environmental stress resistance. The method comprises inserting

```

CC cDNA from a library originating in a salt-resistant organism into a host  
CC cell, culturing the transformants under conditions in which the  
CC untransformed host does not grow well, and selecting for viable clones.  
CC The method is useful for obtaining DNA encoding environmental stress  
CC resistance factors. The DNA encoding proteins conferring environmental  
CC stress resistance, can be used in the production of plants resistant to  
CC environmental stress, which can be cultivated in unfavourable  
CC environments such as deserts, salt damaged ground, cold regions and the  
CC oceans. They can be used for increasing the area of land covered by green  
CC plants, and desert greening and afforestation, in order to counter the  
CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
CC primers AAF74219 and AAF74220 are used in an example illustrating the  
CC method of the invention.

XX Sequence 334 AA:

S0

Query Match 27.9% Score 630.5; DB 22; Length 334;

Best Local Similarity 43.4%; Pred. No. 3.5e-47;

Matches 141; Conservative 67; Mismatches 104; Indels 13; Gaps 5;

OY 132 EGEVTELSPEEASTTGGYAKSISHVITISLKTWGTNQLKIDSSIDALIREKYAVGVDT 191  
||| | :  
Db 2 egevevgidr-pavtgaasktgk---ltlkttemetvydligakmlealgkexvsgdvi 57  
OY 192 YIENSGAVKRVGRCDSPATFYDLAEF--YVPIPKGEVHKKEIYQDVTLLHLDANAAQ 249  
| :  
Db 58 aidkassgkltkgrfsrtdydamgpgvkfvqcdgelgkrkevncvshheidvlnsr 117  
OY 250 POGGODILSLNGOMMKPRKTEITKROEINKVVRNYIDEGIAELVPGVLEIDEVHMIDI 309  
|| :  
Db 118 tqg---flal----ftgtgelraevregidtkvaewreegkaelvpvlfidevhlndi 170  
OY 310 ECFSTLNPALESPLPIVILATNNGICNVKGTDMTSPHGIVDLDRLVYIIRTETGPT 369  
||| : ||| :  
Db 171 ecfstlnralenemapiivvalnrgitlirgtnykpghipidldrlllittgpyckde 230  
OY 370 MIOITAIRAOVEIDIMDEESLAVGEIGQOOTSRLHAIOILSPASVSKTNGREKICKADL 429  
:  
Db 231 lrkldircegedvemaekallthigaelsityahlltaaalacgkkrkglvetedi 290  
OY 430 EVVSGIYLDAKSSARLLDEQOERYI 454  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 291 sraynlfidvkrstqylleyqnygm 315

Search completed: November 13, 2001, 07:01:53  
Job time: 359 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 09:56:10 ; Search time 100.59 Seconds

(without alignments)  
4304.861 Million cell updates/sec

Title: US-09-589-510-3

Perfect score: 1912

Sequence: 1 accacgcgcgcgaattt.....aaaaaaaaaaaaaaaaaaaaa 1912

Scoring table: IDENTITY\_NIC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	612.4	32.0	1750	2 US-08-933-750C-54	Sequence 54, Appl
2	612.4	32.0	1750	3 US-09-234-613-54	Sequence 54, Appl
3	112.6	5.9	703	4 US-08-998-416-1005	Sequence 1005, Ap
4	55.8	2.9	7218	1 US-08-332-463-14	Sequence 14, Appl
5	49.8	2.6	1474	4 US-08-821-994-64	Sequence 64, Appl
6	49.4	2.6	4931	4 US-08-726-320-2	Sequence 2, Appl
7	49.4	2.6	4931	4 US-09-208-716-2	Sequence 7, Appl
8	49.4	2.6	966	1 US-08-514-014-7	Sequence 7, Appl
9	49.4	2.6	966	2 US-08-833-823-7	Sequence 7, Appl
10	48.8	2.6	347	1 US-08-104-072B-2	Sequence 2, Appl
11	48.8	2.6	1117	4 US-09-247-373B-33	Sequence 33, Appl
12	48.8	2.6	3777	3 US-09-121-321-15	Sequence 15, Appl
13	48.8	2.6	3777	3 US-08-933-803A-15	Sequence 15, Appl
14	48.8	2.6	774	4 US-08-956-307B-11	Sequence 11, Appl
15	48.8	2.5	778	4 US-08-821-994-63	Sequence 63, Appl
16	47.8	2.5	1441	4 US-08-924-747-25	Sequence 25, Appl
17	47.2	2.5	991	3 US-08-924-747-25	Sequence 25, Appl
18	47.2	2.5	991	4 US-09-247-373B-25	Sequence 25, Appl
19	47.2	2.5	991	4 US-09-296-715-25	Sequence 25, Appl
20	47.2	2.5	1215	1 US-08-592-214A-1	Sequence 1, Appl
21	47.2	2.5	1215	1 US-09-149-976-1	Sequence 1, Appl
22	47.2	2.5	2760	1 US-08-101-593-1	Sequence 1, Appl
23	46.8	2.4	458	1 US-08-524-757-1	Sequence 1, Appl
24	46.8	2.4	1046	1 US-08-361-467B-4	Sequence 4, Appl
25	46.8	2.4	1046	1 US-08-484-332C-4	Sequence 4, Appl
26	46.8	2.4	6379	1 US-08-499-215-1	Sequence 1, Appl
27	46.6	2.4	940	2 US-08-471-717-1	Sequence 1, Appl

28	46.4	2.4	509	4 US-09-030-607-202	Sequence 202, App
29	46.4	2.4	1512	2 US-08-909-965C-8	Sequence 8, Appl
30	46.4	2.4	3527	2 US-08-909-965C-7	Sequence 7, Appl
31	46.4	2.4	1315	4 US-09-721-822A-10	Sequence 10, Appl
32	46.4	2.4	1364	1 US-08-265-087-3	Sequence 3, Appl
33	46.4	2.4	1364	1 US-08-621-493-3	Sequence 3, Appl
34	46.4	2.4	1364	2 US-08-965-688-3	Sequence 3, Appl
35	46.4	2.4	1364	4 US-09-260-173-3	Sequence 3, Appl
36	45.8	2.4	857	3 US-08-460-040-1	Sequence 1, Appl
37	45.8	2.4	882	2 US-08-909-965C-9	Sequence 9, Appl
38	45.8	2.4	1172	1 US-07-945-288-9	Sequence 9, Appl
39	45.8	2.4	1172	1 US-08-462-831-9	Sequence 9, Appl
40	45.8	2.4	1172	1 US-08-461-809-9	Sequence 9, Appl
41	45.8	2.4	1172	1 US-08-461-441-9	Sequence 9, Appl
42	45.8	2.4	1172	5 PCT-US93-08518-9	Sequence 9, Appl
43	45.6	2.4	688	6 5498694-3	Patent No. 5498694
44	45.6	2.4	2209	1 US-08-514-014-1	Sequence 1, Appl
45	45.6	2.4	2209	2 US-08-833-823-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1

US-08-933-750C-54

Sequence 54, Application US/08933750C

Patent No. 5932442

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Shah, Purvi

APPLICANT: Au-Yang, Janice

APPLICANT: Yue, Henry

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750C

FILING DATE: September 23, 1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

SEQUENCE CHARACTERISTICS:

LENGTH: 1750 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: HMCINOT01

CLONE: 9476  
US-08-933-750C-54

Query Match 32.0%; Score 612.4; DB 2; Length 1750;  
Best local similarity 66.0%; Pred. No. 7.4e-134;  
Matches 902; Conservative 0; Mismatches 461; Indels 3; Gaps 1;

QY 94 atgagatcgaagagtgatgacgtcgaagaaacgacgtccacccacccac 153  
DB 82 ATGAAATTTGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 141  
QY 154 atcaagagacgtcgcctcgcacgcaatggagatgagcattgctgagcggcggtgtg 213  
DB 142 GTGAAGAGGGGTGGGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 201  
QY 214 ggcacagcgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 273  
DB 202 GGCACAG 261  
QY 274 atgagcg 333  
DB 262 ATGAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 321  
QY 334 ctgcgacatgacacgagagcgcgcgcgcgcgcgcgcgcgcgcgcgcg 393  
DB 322 CTGAGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 381  
QY 394 gtagacatcccgagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 453  
DB 382 GTTACTGCACTGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441  
QY 454 ggttgcgtataaagagaacaaagaggttataagagagaggttactgac 513  
DB 442 GGGCTCGGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501  
QY 514 gaagagcgctgagatgacacgtgagatgacacgaagacgttgcacatc 573  
DB 502 TGTAG 561  
QY 574 ttaagacgttataagagagacacgaagacgttgcacatcgttgcacat 633  
DB 562 CTCGAAAG 621  
QY 634 atcaagagaagcgagcgagcgagcgagcgagcgagcgagcgagcgag 693  
DB 622 CAGAAAG 681  
QY 694 aagaagctgtatgagatgacgttgcacatcgttgcacatcgttgcac 753  
DB 682 AAG 741  
QY 754 gtlcctatcccaagagtgagtgacatgagaaagaaatltgagagatg 813  
DB 742 GTCCCTTGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 801  
QY 814 catgacgttgcacgacgaatgacgaacgaatgacgaacgaatgacga 873  
DB 802 CATGACTTGAGATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 861  
QY 874 ggcacagatgataaagcgaagacgtgagacgaacgaacgaacgaac 933  
DB 862 GGCACAG 921  
QY 934 aaggtgtgataatagatatacgaagagagatgacagagctgtgacgtg 993  
DB 922 AAGGTGTGAACAAGTACATGACAGAGAGAGAGAGAGAGAGAGAG 981  
QY 994 attgagagcgctgacatgtagatgtagatgtagatgtagatgtagat 1053  
DB 982 GTTGTAG 1041  
QY 1054 agcccaatcacaacatgtagatgtagatgtagatgtagatgtagat 1113

DB 1042 TCTTATACGCTCCCATCGATCATCTTGGATCCAGAGAGAGAGAGAG 1101  
QY 1114 act--gatalgacagatccacatgtagatagcggtagatcttagatagcgg 1170  
DB 1102 ACTGAGAGATCAATCCATCCCTACAGAGAGAGAGAGAGAGAGAGAG 1161  
QY 1171 attgagagagatgtagatgtagatgtagatgtagatgtagatgtagat 1230  
DB 1162 ATCCGAG 1221  
QY 1231 gtagagagatgtagatgtagatgtagatgtagatgtagatgtagat 1290  
DB 1222 ACGAAG 1281  
QY 1291 acatctttagagatgtagatgtagatgtagatgtagatgtagatgtag 1350  
DB 1282 ACCACAG 1341  
QY 1351 ggaag 1410  
DB 1342 GGAAG 1401  
QY 1411 aaatctcgctcgctcgctcgctcgctcgctcgctcgctcgctcgct 1456  
DB 1402 AAGTCTCGCCCAAAATCTGCTGACAGAGAGAGAGAGAGAGAGAG 1447

# RESULT 2

US-09-234-613-54  
Sequence 54, Application US/09234613  
Patent No. 6132973

## GENERAL INFORMATION:

APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA

COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSO for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1750 base pairs







RESULT 8  
US-08-514-014-7  
; Sequence 7, Application US/08514014

```

1  TITLE OF INVENTION: ENCODED THEREBY
2  NUMBER OF SEQUENCES: 12
3  CORRESPONDENCE ADDRESS:
4  ADDRESS:
5  STREET: 87 Cambridgepark Drive
6  CITY: Cambridge
7  STATE: Massachusetts
8  COUNTRY: USA
9  ZIP: 02140
10
11  COMPUTER READABLE FORM:
12
13  MEDIUM TYPE: Floppy disk
14  COMPUTER: IBM PC compatible
15  OPERATING SYSTEM: PC-DOS/MS-DOS
16  SOFTWARE: Patent In Release #1.0, Version #1.25
17
18  CURRENT APPLICATION DATA:
19  APPLICATION NUMBER: US/08/833,823
20  FILING DATE: 10-APR-1997
21
22  CLASSIFICATION: 530
23
24  PRIOR APPLICATION DATA:
25  APPLICATION NUMBER: 08/514,014
26  FILING DATE: 11-AUG-1995
27
28  ATTORNEY/AGENT INFORMATION:
29  NAME: Brown, Scott A.
30
31  REGISTRATION NUMBER: 32,724
32  REFERENCE/DOCKET NUMBER: G16000
33  TELECOMMUNICATION INFORMATION:
34  TELEPHONE: (617) 498-8224
35  TELEFAX: (617) 876-5851
36
37  INFORMATION FOR SEQ ID NO: 7:
38  SEQUENCE CHARACTERISTICS:
39  LENGTH: 966 base pairs
40  TYPE: nucleic acid
41  STRANDEDNESS: double
42  TOPOLOGY: linear
43  MOLECULE TYPE: cDNA
44  HYPOTHETICAL: NO
45  FEATURE:
46
47  NAME/KEY: CDS
48  LOCATION: 67..348
49
50  OS-08-833-823-7

```

Query Match	2.6%	Score 49	DB 2	Length 966
Best Local Similarity	55.8%	Pred. No. 0.017	Mismatches 91	Conservative 1
Matches	91	Conservative	1	Mismatches 71
				Indels 0
				Gaps 0
QY	1750	cggagttgattggccgaaccctactcttgcataccatgataagatgaattcttaac	1809	
Db	781	CTCTGCTATVTGATCTTTATGCTATATTACATCTCTGTGCTAGCTGGAACATTTGACAT	840	
QY	1810	gaatgcgaacttcgcatgagcttatacttctctaaatgctccataaagcacaagcaatglttc	1869	
Db	841	TATTACTGGAGTCAGGCCCTTATAAGTCAAAAGCACCTATGTGCTGAAGCACTTCTCTCA	900	
QY	1870	tacaacmtvtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	1912	
Db	901	AACATTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	943	

RESULT 1072B-2  
US-08-104-072B-2  
Sequence 2, Application US/08104072B  
Patent No. 5639948  
GENERAL INFORMATION:  
APPLICANT: Michiels, Frank  
APPLICANT: Morioaka, Sinji  
APPLICANT: Scheirlinck, Trees  
APPLICANT: Komari, Toshiko  
TITLE OF INVENTION: Stramen-specific Promoters from Rice  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 North 5639948west Center

```

1  APPLICANT: O'KEEFE, DANIEL
2  TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
3  FILE REFERENCE: CL-1108-A
4  CURRENT APPLICATION NUMBER: US/09/247,373B
5  CURRENT FILING DATE: 1999-02-10
6  PRIOR APPLICATION NUMBER: 08/924,747
7  PRIOR FILING DATE: 1997-09-05
8  NUMBER OF SEQ ID NOS: 56
9  SOFTWARE: Microsoft Office 97
10 SEQ ID NO 33
11
12 LENGTH: 1117
13
14 TYPE: DNA
15
16 ORGANISM: SOYBEAN
17
18 FEATURE:
19
20 NAME/KEY: unsure
21 LOCATION: (1101)
22 OTHER INFORMATION: M-A OR C
23
24 NAME/KEY: unsure
25 LOCATION: (1104)
26 OTHER INFORMATION: M-A OR C
27
28 NAME/KEY: unsure
29 LOCATION: (1116)
30 OTHER INFORMATION: N-G OR A OR T OR C
31
32 US-09-247-373B-33

```

```

Query Match          2.6%; Score 48.8; DB 4; Length 117;
Best Local Similarity 73.8%; Pred.No.0.02; Mismatches 19; Indels 0; Gaps 0;
Matches 59; Conservative 2;

QY      1833 tatctcgaatggtccataagacatacgaaatggttcttcacaacmblwlaaaaaaaaaaaaaa 1892
        | | | | | | | | | | | | | | | | | | : | | | | | | | | | |
Db       960 ttatatatgatggcaataaactcatattcattctttaaaaaaaaaaaaaaaaaaaaaa 1019

QY      1893 aaaaaaaaaaaaaaaaaaaa 1912
        ||| ||||| ||||| |||||
Db       1020 aaaaaaaaaaaaaaaaaaaa 1039

RESULT 12
US-09-121-321-15
; Sequence 15, Application US/09121321
; Patent No. 6090783
; GENERAL INFORMATION:
; APPLICANT: Saiga, Akiniko
; APPLICANT: Orita, Satoshi
; APPLICANT: Igarashi, Hisanaga
; APPLICANT: Okumura, Kouichi
; APPLICANT: Sakaguchi, Gaku
; TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF
; TITLE OF INVENTION: GENE EXPRESSION AND NOVEL PROTEIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/121,321
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/933,803
; FILING DATE: 19-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.

```

SEQUENCE CHARACTERISTICS:

S-08-956-307B-12



Wed Nov 14 08:34:11 2001

us-09-589-510-3.std.mri

---

Page 10



Copyright (c) 1993-2000 CompuGen Ltd.

```
-MODEL=frame+p2n.model -DEV=xlp
-runcat /usr08589510/runcat 13112001_065634_13459/app_query.fasta_1.518
```

Search information block:  
query: US-09-589-510-4

```
Query length: 351203
Database: Issued_Patents_NA: *
Database sequences: 351203
Database length: 113238999
Search time (sec): 93.430000
```

score\_list:

```

/cgn2_6/prodata/2/ina/5B_COMB.seq:US-09-433-262-1 + 95.50 131.98 10.60 1407
/cgn2_6/prodata/2/ina/6A_COMB.seq:US-08-859-694-1 + 95.50 131.55 11.18 1457
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-309-341-3 - 95.00 154.54 0.5863 2002
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-608-267-3 - 95.00 154.54 0.5863 2002
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-608-452-3 - 95.00 154.54 0.5863 2002
seq_name: /cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-933-750C-54

seq_documentation_block:
; Sequence 54, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Yang, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1750 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 9476
; US-08-933-750C-54

alignment_scores:
Quality: 1721.50 Length: 455
Ratio: 4.109 Gaps: 1
Percent similarity: 92.088 Percent identity: 72.527

alignment_block:
US-09-589-510-4 x US-08-933-750C-54 ..
Align seg 1/1 to: US-08-933-750C-54 from: 1 to: 1750
1 MetAgtTlGtGtGtAValGInSerThrSerTylsGInatGtTleatTh 17
||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1317 gAlaIeuGluSerProIeuSerProIleValIleuAlaItrhAsnArgC 334
1032 CCCCCTGAGCTCTCTATCTCTGCTCCATCCGTCATCTTGCATCCAAACCGAG 1081
334 LyIleCysAsnValaIaRgIYthr...AspMetThrSerProHisGlyIle 349
1082 GCACCTGTGTCATCAGAGGACGACATCAGACATCCATCCCTACGGCATC 1131
350 ProValAspIleuLeuAspArgIleuValIleIleIaYtrhGluYtrG1 366
1132 CCTCTTGACCTTCTGCGACCCGAGTGAATGATGAATCCGGACCATGCTGTATAC 1181
366 YProThGluMetIleGlnIleuAlaIleArgIaGlnValGluG1 383
1182 TCCACAGGAATGAACACAGATCATTAATAATCCGTGCCAGACGGAAGAA 1231
383 LeAspMetAspGluGluSerIleuAlaIaYtrIleuGlyGluIleGlyGln 399
1232 TCAACATCAGTGTAGGAGGACATGAAACCACTGGGGAGATTGGACCAAC 1281
400 ThrSerIleuAtrhIleAlaIleGlnIleuIleSerProAlaSerValaIse 416
1282 ACCACACTAGGTACTCAGTGCAGCTGCTGCACCCCGCCAACTTCTTGC 1331
416 rLyStrhAsnGlyAaGgLuLyIleCysIlyAsnIaAspLeuGluGlnValS 433
1332 TAAATCAACAGCGGAGACACACATTGAGAAAGACATGTGCAAGACATCA 1381
433 eRgIYleuTYrIleuAspAlaIySerserAlaargIeuGluGlnGluG1 449
1382 GTGACATTTCTATGATGTCACAAAGTCCTCCGCAAAATCTGCTGACCG 1431
450 GlnGluAaTYrIle 454
1432 CAGGATTAAGTACATG 1446

seq_name: /cogn2_6/prodata/2/ina/6A.COMB.seq:US-09-234-613-54
seq_documentation_block:
? Sequence 54, Application US/09234613
? Patient No. 6132973
? GENERAL INFORMATION:
? APPLICANT: Lal, Preeti
? APPLICANT: Hillman, Jennifer L.
? APPLICANT: Bandman, Olga
? APPLICANT: Shah, Purvi
? APPLICANT: Au-Young, Tanice
? APPLICANT: Yue, Henry
? APPLICANT: Cuesqier, Karl J.
? TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
? NUMBER OF SEQUENCES: 98
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Incyte Pharmaceuticals, Inc.
? STREET: 3174 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/234,613
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/933,750
? FILING DATE: September 23, 1997
? ATTORNEY/AGENT INFORMATION:

```

Wed Nov 14 08:34:16 2001

us-09-589-510-4.std.in1

NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1750 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HMC1N0701  
CLONE: 9476  
US-09-234-613-54

alignment scores:  
Quality: 1721.50 Length: 455  
Ratio: 4.109 Gaps: 1  
Percent Similarity: 92.088 Percent Identity: 72.527

alignment block:  
US-09-589-510-4 x US-09-234-613-54 ..

Align seg 1/1 to: US-09-234-613-54 from: 1 to: 1750

1 MetArgIleGluGluValAlaGlnSerThrSerLysGlnArgIleAlaThr 17  
82 ATGAGCATTTGAGGAGGTGAAGACACACAGACGAGGCGATGCCCTC 131  
17 ThrHisThrIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIle 34  
132 CCACAGCCACGTGAAGGGCTGGGGCTGGACGAGAGCGGCTTGGCAGC 181  
34 LeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaGly 50  
182 AGCGGCGCTCAGGCGCTTGGGCGAGGAGAACCGCGGAGGCGATGGC 231  
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyAlaValIle 67  
232 GTCTAGTATGAAATTAATGCAAGCAAGAAATGCGTGGAGAGCTGCTT 281  
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84  
282 GTTGGCAGGACCTCTGGAGACTGGCAAGACACTCTGCTGCTGCTATTG 331  
84 IagGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100  
332 CTCAGAGACCTGGTATGATAGCTCCCTCTGCCCCAATGGTGGGAGTGA 381  
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117  
382 GTTACTACTAGATCAACAGACAGACAGAGTCTGATGGAGAACTCCG 431  
117 GArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGlnGly 134  
432 CAGGGCCATGGGCTCGCAATTAAGAGACCAAGAGATTATGAGAGTG 481  
134 LuValThrGluLeuSerProGluGluAlaGlnSerThrThrGlyGlyTyr 150  
482 AAGTTCACAGACTAATCTCGTGTGAGACAGAGAAATCCCATGGGAGATAT 531  
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGly 167  
532 GGCAGAAACATTAGCCATGTATGATAGACCTCAACAGGCAAGAGAAC 581  
167 PLYGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184  
582 CAAAGCTGTAAGTGAACCCAGCATTTTGAAGATTGTCAGAAAGGC 631

184 yValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGluValAla 200  
632 GAGTAAAGCTGAGATGATGATTAATGATTAAGCAACAGAGGGGCGCTG 681  
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
682 AAGAGGAGGAGGAGGTGATACCTATGCGACAGAAATTCGACTTGAAGC 731  
217 agLysGluTyrValProIleProLysGlyGluValHisLysLysGluI 234  
732 TGAAGATATGCTCCCTTGCAGAAAGGAGATGTCAGCAAGAAAGAGAA 781  
234 LeuAlaGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250  
782 TCATCCAAAGATGTACCTTCCATGACTTGGATGATGGCTTAATCGGGGCC 831  
251 GlnGlyLysGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267  
832 CAGGGGAGCAAGATATCTGTCATGATGAGGCGAGCTAGCTAATGACCA 881  
267 GlyThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284  
882 GAGACAGAAATCAACAGCAACTTGGAGGAGATTAATAGTGTGTA 931  
284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300  
932 ACAAGTACATGACGACGAGCATTTGCTGAGTGGTCCGGGTGCTGCTT 981  
301 IleAspGluValHisMetLeuAspIleGlnCysPheSerTyrLeuAsnAr 317  
982 GTTGAGAGAGTCCACATGCTGACATTCAGTCTGCTACCTGACGACG 1031  
317 gaLysGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334  
1032 CGCCCTGAGGTCTTATGCTCCATGCTGCTGCTGCTGCTGCTGCTGCT 1081  
334 yLysGluAsnValArgGlyThr...AspMetThrSerProHisGlyIle 349  
1082 GCACCTGTGTCATCAGAGGACGACGAGCATCATCCCTCAGGCGATC 1131  
350 ProValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrG 366  
1132 CCTCTGACCTTCTGACCGAGATGATGATATCGGACCATTCCTGTATAC 1181  
366 yProThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGlu 383  
1182 TCACAGGAAATGAACAGATCATTAATTCGCGCCAGAGGAGGAA 1231  
383 LeAspMetAspGluGlnSerLeuAlaTyrLeuGlyGluIleGlyGlnGln 399  
1232 TCACATCAGTACGAGGAGGACCTGACCAACCTGGGGAGATGGCAGCAG 1281  
400 ThrSerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValAlaSe 416  
1282 ACCACACTAGAGTACTCAGTGTGAGTGTGCTGACCCCGGCAACTGTGTC 1331  
416 rLysThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluGluValS 433  
1332 TAAATCAACAGGAGGAGACAGCATTTGAGAAAGACATCTGCAGAAATGCA 1381  
433 ereGlyLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGln 449  
1382 GTGAACCTTCTATGATGCCAAGTCTCGCCAAATCTGCTGACACAG 1431  
450 GlnGluArgTyrIle 454  
1432 CAGGATTAAGTACATG 1446  
seq.name: /cgn2\_6/prodata/2/lna/6R\_COMB.seq: US-08-998-416-1005  
seq.documentation\_block:  
; Sequence 1005, Application US/08998416  
; Patent No. 6239264

## GENERAL INFORMATION:

APPLICANT: Philippson, Peter  
 APPLICANT: Pohlmann, Rainer  
 APPLICANT: Steiner, Sabine  
 APPLICANT: Mohr, Christine  
 APPLICANT: Wendland, Jürgen  
 APPLICANT: Knechtle, Philipp  
 APPLICANT: Redischung, Corinne  
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSPYII  
 NUMBER OF SEQUENCES: AND USES THEREOF  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6239264rtis Corporation  
 STREET: 3054 Cornwalis Road  
 CITY: Research Triangle Park  
 STATE: No. 6239264th Carolina  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/998,416  
 FILING DATE: 24-DEC-1997  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: CH 0016/97  
 FILING DATE: 31-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGJ1976  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 1005:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 703 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: PAG1617RP  
 US-08-998-416-1005

## alignment\_scores:

Quality: 312.00 Length: 126  
 Ratio: 3.120 Gaps: 1  
 Percent Similarity: 79.365 Percent Identity: 50.794

## alignment\_block:

US-09-589-510-4 x US-08-998-416-1005/rev ..

Align seg 1/1 to reverse of: US-08-998-416-1005 from: 1 to: 703

```

1 MetATGllGluGluValGlnSerThrSerLysLys.....GlnATGll 15
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
378 ATTCAACACAGAGAGGCGCATGCTCTGTAAGTCTTGTGCGCAAT 329
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
15 eAlATpThHstHstHstLysGlyLeuGlyLeuAspAlaAsnGlyMetA 32
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
328 TCGACACACTTCGACACTTGTGGCTGGGACTGACGACGACCTACAGC 279
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
32 lAlleAlaLeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAla 48
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
278 CTAAAGCCAGACTCGACGAGGATGTGGCCAGATTACAGGCGCGGAGCGCC 229
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
49 lAlaGlyLeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAl 65
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

```

228 GCAGCGGTGATTCTCAAAATGTGCAAGACGTAACGATTCCGGGGGGC 179
65 aValleuAlaGlyPropRoAlaThrGlyLysThrAlaLeuAlaLeuG 82
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
178 CGTGTGTTGCTGCTGCTCGCGCTGACCGCAAAAGCGCGTGGCAGATG 129
82 lYlleAlaGlnGluLeuGlySerLysValAlaPheCysProMetAlaGly 98
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
128 GCTTGTGCACTGCTGCGCGAGATGTGCTTACACGCAATGCGCAGG 79
99 SerGluValIyrSerSerGluValLysLysThrGluValLeuMetGluAs 115
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
78 TCAAGATATTCTGCTGTTGGAGCTATCCAGACGAGCGCTTGACTACAGC 29
115 nPheAlaArgAlaAlaGlyLeuArgIle 124
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
28 GTTCCGAAAGTCGATGTGTCAGATC 1

```

## seq\_name: /cgn2\_6/ptodata/2/lna/5b\_comb.seq:us-09-103-840A-1

## seq\_documentation\_block:

; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4411529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37Rv  
 US-09-103-840A-1

## alignment\_scores:

Quality: 131.00 Length: 504  
 Ratio: 0.582 Gaps: 29  
 Percent Similarity: 44.643 Percent Identity: 22.024

## alignment\_block:

US-09-589-510-4 x US-09-103-840A-1/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-1 from: 1 to: 4411529

```

1 MetATGllGluGluValGlnSerThrSerLysLysGlnArgIleAlaThr 17
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
4052623GTACGCGCTGCACTGTTCACAGCGCTCAGCGCAAAAGCGGAAGTCA 4052574
17 rHstHstHstHstLysGly.....LeuGlyLeuAspA 28
   ||| ::::: ::::: ::::: ::::: ::::: :::::
4052573CCACAGGTGTCACACCAAGCATCTGCGGACACTGCTGCTACGCTG 4052524
28 lAsnGlyMetAlaAlaAlaLeuAlaAlaGlyPheValGlyGlnAla... 43
   ||| ::::: ::::: ::::: ::::: ::::: :::::
4052523CT.....GCGGCTGCTGTTGCTGGTGGGCGTCTCGATGTTCTC 4052483
44 .....AlaAlaArgGluAlaAlaGlyPheValAlaValAspMetIleArg 58
   ||| ::::: ::::: ::::: ::::: ::::: :::::
4052482CCCGCATGCAAGCGCGCGCGGATGGCTTGGGATTCGCAAGTCACGCG 4052433
58 lNlysLysMet.....Ala 62
   ||| ::::: ::::: ::::: ::::: :::::
4052432CCAGACACTGAGCAAGACATGCCCAAGACACACTTCGCGCAGATCGCA 4052383

```



```

41 ygl.....AlaAlaAlaArgGluAlaAlaGlyLeuAlaValA 54
   |||..... |||..... |||.....
2875506TTCGCGAGGACGACCTGCTGCGCGCGCGTTCGCGCGCGCT... 28755861
54 spmetIleArgGlnLysLysMetLeuAclyArgAlaValLeuLeuAcl 70
   |||..... |||..... |||..... |||.....
2875660.....GTCGAGGCGTGGGGGTGGCGTGGTGTCTCTACAG 2875822
70 yProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleAlaGlnGlu 87
   |||..... |||..... |||..... |||.....
2875821TCCCGCGCGCGCGAGCAAGCAAGCGCGCGGTGATCTGCAGGCCA 2875772
87 euGlySerLysValProPhcCysProMetValGlySerGluValYrSer 103
   |||..... |||..... |||..... |||.....
2875771CCCGCGCGCGG.....TTCGAGCGCCTGTCGGCATGTGCG 2875737
104 SerGluValLysLysThrGluValLeuMetGluAsnPhcArgAlaAla 120
   |||..... |||..... |||..... |||.....
2875736GCGCGCGTCAAGAGAGAGCGCGGTGATCTGCAGAACATTCGCAAGACGCT 2875687
120 eGlyLeuAlaArgIleLysGluAsnLysGluValYrGluGluValYrG 137
   |||..... |||..... |||..... |||.....
2875686.....CTTCAGCGCGCGAGCAGACCG 2875667
137 IuLeuSerProGluGlu.....AlaGluSerThrThrGly 149
   |||..... |||..... |||..... |||.....
2875666TGTGCTGATCGACGAGGTGACCGCATTTTCCAGAGCCACACAGACCGCC 2875617
150 TyrAlaLysSerIleSerHisValIleLeuSerLeuLysThrVallyG 166
   |||..... |||..... |||..... |||.....
2875616CTCTCTGCGCGGTGAGACACCGGGTGGCGTGG.....GTGCGCGC 2875573
166 yThrLysGlnLeuLysLysLeuAspSerIleYrAspAlaLeuLysG 183
   |||..... |||..... |||..... |||.....
2875572GACACCAACGAGAACCCGCTGCTTTCGTTGCGTGCACCGCTGTCGCGCT 2875523
183 IuLysValAlaValGlyAspValIleYrIleGluAlaAspSerGlyAla 199
   |||..... |||..... |||..... |||.....
2875522CGCTGATCTGACGCTGCGCGCGCTGACGCGCGAGACACCGCGCGGTA 2875473
200 ValLysArg.....ValGlyArgCysAspSerPh 209
   |||..... |||..... |||..... |||.....
2875472GTACCAACGCGCATGATGACCGCGCGCGCGTGGAGGCGAGCGCGGT 2875423
209 eAlaThrGluYrAspLeuGlnAlaGluGluYrValProIleProLysG 226
   |||..... |||..... |||..... |||.....
2875422ACCGCGCTGAG.....GCCGTGACCTGCTGTGCACTCCGCGCGC 2875382
226 IyGluValHisLysLysGluIleValGlnAspValThrLeuHisAsp 242
   |||..... |||..... |||..... |||.....
2875381GCGAGCGCGCGCGC.....GCCCTGACCGCG 2875356
243 LeuAspAlaAlaAsnAlaGlnProGluGlyGlyLysAspIleLeuSer 259
   |||..... |||..... |||..... |||.....
2875355CTGGAGAGTACGCGCGAGCGAGCGCGCGCGCGCGTGTGACGCGT 2875306
259 uMetGlyLysMetLysProArgLysThrGlnIleThrGluLysLeuA 276
   |||..... |||..... |||..... |||.....
2875305G.....CAACACATCG 2875295
276 rGlnGluIleAsnLysValAlaAsnArgTyrIleAspGluGlyIleAla 292
   |||..... |||..... |||..... |||.....
2875294AGCGGTGAGACAGCGCGGTGCGTACAGACGTCAGCGCGACGAG 2875245
293 GluLeuValProGlyValLeuPheIleAspGluValHisMetLeuAsp 309
   |||..... |||..... |||..... |||.....
2875244CAGCTACGACCTGTCAGCGCGCTTTCATCAATCGTGGCGGTTCGACGT 2875195
309 eGlu.....CysPheSerTyrLeuAsnArgAlaLeuGluSerProLeuSerP 325
   |||..... |||..... |||..... |||.....
2875194CGATGCGCGCTGACATTTGGCCCGCATGCG..... 2875161
325 rIleValIleLeuAlaThrAsnArgIlyIleCysAsnValArgGlyThr 341

```

```

2875160..... |||..... |||.....
342 AspMetThrSerProHisGlyIleProValAspLeuLeuAspArgLeuValA 358
   |||..... |||..... |||..... |||.....
2875148GAC.....CCGCGTTATGCGCGCGCGCTGAT 2875120
358 IleIleArgThrGluThrTyrGlyProThr.....GluMetIleGlnI 373
   |||..... |||..... |||..... |||.....
2875119GATTTCTGCCAGAGACATCGCATGCGCGCGCGCGCGTGCAGG 2875070
373 LeuAlaIleArgAlaGlnValGluGluIleAspMetAspIuLeuSer 389
   |||..... |||..... |||..... |||.....
2875069TGGCGCGCGCGCGCGCACAG.....ACG 2875047
390 LeuAlaTyrLeuGlyGluIleGlyGlnGlnThrSerLeuArgHisAla 405
   |||..... |||..... |||..... |||.....
2875046GTGCGCTGATCGCAGTGGCCGAGACCCAGCTGACGCTGCGCGACGCC 2874999

seq_name:/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-999-774A-5

seq_documentation_block:
; Sequence 5, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gisp, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: Mcclanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; NUMBER OF INVENTION: Related Reagents
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34, 090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2932 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 266..2683
; US-08-999-774A-5

alignment_scores:

```

Quality: 117.00 Length: 219  
Ratio: 1.045 Gaps: 13  
Percent Similarity: 51.142 Percent Identity: 25.114

## alignment block:

US-09-589-510-4 x US-08-999-774A-5 ..

Align seg 1/1 to: US-08-999-774A-5 from: 1 to: 2932

```

64 ArgAlaValLeuLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAl 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
980 AGAGATTCCTGCTTACGACCTCCGAGACAGAGAAAGCCGTGATGC 1029
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
80 AlaGlyIleAlaGlnGluLeuGlySerLysValProPheCysPromely 97
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1030 TCGAGCTGTAAGTAATGAGACTGAGCCTTC.....TTCTTCTGATCA 1073
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
97 aLgysrGluValTyrSerSerGluValLysLysThrGluValLeuMet 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1074 ATGCTCTGAGATCATGAGCAATGCTGCTGATGCTGAG..... 1114
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
114 GluAsnPheArgArgAlaIleGlyLeuArgIleLysGluAsnLysGlu.. 129
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1115 AGCAACCTTCTGTAAGCCTT.....GAGAGGCTGAGAA 1149
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
130 .....ValTyrGluGlyGluValThrGluLeuSerProG 141
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1150 GAATGCTCTGCTCCATCTTCATTCATTCATTCATTCATTCATTCATTC 1199
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
141 LngLualGluSerThrThrGlyLysThrAlaLysSerIleSerHisVal 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1200 AAAAAA...GAAAACTCATGCGAGGTGAGCGCGCCATTGATCATCAG 1246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
158 IleIleSerLeuLysThrValLysGlyThrLysGluLeuLysLeuAsp 174
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1247 TTGTTGACCTC.....ATGATGCGCTTAAGCAGAG..... 1279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 rSerIleTyrAspAlaLeuIleLysGluLysValAlaValGlyAspAl 191
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1280 .....GCACATGTGA 1289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
191 LeTyrIleGluAlaAsnSer.....GlyAlaValLys 201
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1290 TTGTTATGCGACGACCAACAGACCAACGATTCGACGCTCTACGG 1339
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
202 ArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAla 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1340 CGATTGTGCTGC.....TTTGACAGGAGGTAGA 1368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 uGluTyrValProIleProLysGly.....GluValHis 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1369 TATTGCAATTCCTGATGCTACAGGAGCCTTAGAGATTCCTCAGATCAT 1418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
230 ..LysLysLysGluIleValGlnAspValThrLeuHisAspLeuAspAl 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1419 CCAAGACATGAACTGCGAGATGATG.....GACCTGGAACAG 1459
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
246 AlaAsnAlaGlnProGlnGly.....GlyGlnAspIleLeuSerLeu 260
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1460 GTAGCCATGAGACTCAGCGGCGATGGTGTGCTGACTTAGAGCCCTGTG 1509
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
260 tGlyGln 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1510 CTCAGAG 1516
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: /cgn2\_6/prodata/2/lna/5B\_COMB.seq:US-09-446-504-4

## seq\_documentation\_block:

; Sequence 4, Application US/09446504  
; Patient No. 6218150  
; GENERAL INFORMATION:  
; APPLICANT: UEMORI, Takashi  
; APPLICANT: SATO, Yoshimi

```

APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazuo
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/446,504
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 984
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-446-504-4

```

alignment\_scores:  
Quality: 111.50 Length: 367  
Ratio: 0.664 Gaps: 18  
Percent Similarity: 45.777 Percent Identity: 21.526

## alignment block:

US-09-589-510-4 x US-09-446-504-4 ..

Align seg 1/1 to: US-09-446-504-4 from: 1 to: 984

```

54 AspMetIleArgGlnLysLysMetAlaGlyArgAla..... 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
76 GACATTGTGAGACAGACGACATAGTGAAGAGCTCAACGACTAGCTCAA 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 .....ValLeuLeuAlaGlyProProAlaThrGlyL 76
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
126 AACTGATCAATGCCCCACTACTCTTCGAGAGCCCGCTGCTGCGAA 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
76 ySthrAlaLeuAlaLeuGlyIleAlaGlnGluLeuGlySerLysValPro 92
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
176 AGACTACAGCGCTTTGGCCCTTGCAAGAGAGCTT..... 210
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
93 PheCysPromelyAlaGlySer.....GluValTyr 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 .....TTGCGCGAAAACTGAGGACATTAACCTTCCTGAGTTGAA 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 rSerSerGluValLysLysThrGluValLeuMetGlu.....Asp 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 TGCTTCAGATGAAGAGGATATACGTATATAGACGAAAGTTAAGAGAT 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
116 heArgArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGlu 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
299 TTGCGAGAACAAAGCCTTATAGAGAGCAAGCTTCAAGATATATTCCTT 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
133 GlyGluValThrGluLeuSerProGluGluAlaGluSer.....Th 146
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
349 GATGAGCGCGACCTTTAACTCAAGATGCCCAACAGCCTTAGAAGAAC 398
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
146 rThrGlyGlyTyrAlaLysSerIleSerHisValIleIleSerLeuLysT 163
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
399 CATGAAATGTTCTCGAGTACCTTCCTTATCTTGAGCTGTAATTA 448
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
163 hrValLysGlyThrLysGlnLeuLysLeuAspSerSerIleTyr..... 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
449 CCGCAAGATATTTGAACCATACAGCTAGATGTCAATATTCGCTTC 498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
178 .....AspAlaLeuIleLysGluLysValAlaValGlyAspAl 190
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
499 AGACTCTCGCGATGAGATATAGCGAAGAGACTAAG..... 537
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

190 111eTyrIleGluAlaAsnSerGlyAlaValLysArgValGlyArgCysA 207
    ||||| ||||| |||||
538 ....TACATTGCCGAAATAGGCGCTTAGAG..... 564
207 spSerPheAlaThrGluTyrAspLeuGluAlaGluGluTyrValProIle 223
    ||||| ||||| |||||
565 .....CTAAGTGAAGAAGGTCTCCAAACCACTACTTTC.....ATA 600
224 ProLysGlyGluValHisLysLysLysGluIleValGlnAspValThrIle 240
    ||||| ||||| |||||
601 GCAGAAAGAGATATGAGAAAGCAATTAACATTCTGCAAGCTGCACAGAG 650
240 uHisAsp.....LeuAspAlaAlaAsnA 248
    ||||| ||||| |||||
651 TCTAGACAGAGATCATCCGACGAAACCTATTTCATGTCGATGACGATGAG 700
248 1aGlnProGlnGlyGlnAsp.....IleLeuSerLeuMetGlyGln 262
    ||||| ||||| |||||
701 CTAGACCTGAGATATAGAGAGATGATGCTTCTCTCTCAAGGCAAC 750
263 MetMetLysProArgLysThrGluIleThrGluLysLeuArgGlnGlu 279
    ||||| ||||| |||||
751 TCTTGAAGGCCAGA.....GAAAGCTTAGGAGATGACT 785
279 eAsnLysValAlaAsnArgTyrIleAspGluGlyIleAlaGluLeuValP 296
    ||||| ||||| |||||
786 TCTCAAG.....CAAGGACTTACT..... 804
296 roGlyValLeuPheIleAspGluValHisMetLeuAspIleGlyCysPhe 312
    ||||| ||||| |||||
805 ..GGAGAAGATGACTAGTTCAGATCCACAAA.....GAACTCTTC 843
313 SerTyrLeuAsnArgAlaLeuGlnSerProLeuSerProIleValIleLe 329
    ||||| ||||| |||||
844 .....AACCTGCCAATAGAGGCCCAAGAG.....GTTCTGCT 878
329 uAlaThrAsnArgGlyIleCysAsnValArgGlyThrAspMetThrSerP 346
    ||||| ||||| |||||
879 TGCTGATTAAGATAGAGAGATATTAATCTGACTCGTTGAA..... 918
346 rHisIleGlyLeuProValAspLeuAspArgLeuValIleLeuArgThr 362
    ||||| ||||| |||||
918 ..... 918
363 GluThrTyrGlyProThrGluMetIleGlnIleLeuAlaIleArgAlaG 379
    ||||| ||||| |||||
919 .....GGGCTTAATGAATAATATTCAGCTTGAAGCACCTCTTAGCAC 959
379 n 379
    |
960 G 960
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-08-987-123-1
seq_documentation_block:
; Sequence 1, Application US/08987123C
; Patent No. 6136557
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; FILE OF INVENTION: STREPTOCOCCUS PNEUMONIAE GENE SEQUENCE ftsH
; TITLE REFERENCE: X-11762 Sequence Listing
; Patent No. 6136557
; CURRENT APPLICATION NUMBER: US/08/987,123C
; PRIOR FILING DATE: 1997-12-08
; PRIOR APPLICATION NUMBER: 60/036,281
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1956)
US-08-987-123-1

Alignment_scores:
  Quality: 111.50      Length: 484
  Ratio: 0.474         Gaps: 31
  Percent Similarity: 48.554      Percent Identity: 23.140

alignment_block:
US-09-589-510-4 x US-08-987-123-1

Align seg 1/1 to: US-08-987-123-1 from: 1 to: 1959

39 pheValIGlyGlnAlaAlaArgGluAlaGluLysLeuAlaValAspPhe 55
    ||||| ||||| |||||
556 TTTTCAGATGTTGCTGAGGAGTGAAGAAAACAGAACTAGTTGAAGT 605
55 fLe.....ArgGlnLysMetAlaGlyArg. 64
    ||||| ||||| |||||
606 TGTTCAGTCTTAAAGATCCAAACGATTCACAAACTGGAGCCGTA 655
65 .....AlaValLeuLeuAlaGlyProProAlaThrGlyLysThrAla 78
    ||||| ||||| |||||
656 TTCACAGAGGTGTTCTTGGAGGAGACCTCCGGGACAGGTAAACCTTG 705
79 LeuAlaLeuGlyIleAlaGlnGluLeuGlySerLysValProPheCysPr 95
    ||||| ||||| |||||
706 CTTCCTAAGGAGTGGCTGAGAGAGAGT.....GTTCAATCTTTAG 749
95 oMetValGlySerGlu.....ValTyrSerSerG 105
    ||||| ||||| |||||
750 TATTCAGAGTCTTACCTTTAGAAATGTTTGGCGGTTGAGCTTAGTC 799
105 LuValLysLysThrGluValLeuMetGluAsnPheArgAla..... 119
    ||||| ||||| |||||
800 GTGTGCTCT.....CTTTTGAGATGCAAAAAAGCAGCCA 840
120 .....IleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
    ||||| ||||| |||||
841 GCTATCATCTTTATCGATCTAAATGATGCTGTTGACGTCACGTCGAGT 890
134 uValThrGluLeuSerProGluGluAlaGlnSerThrThrGlyGlyTyrA 151
    ||||| ||||| |||||
891 CGGCTGGGGGAGGTAAAGCAGCTGAACAAC..... 927
151 lAlysSerIleSerHisValIleIleSerLeuLysThrValLysGlyThr 167
    ||||| ||||| |||||
928 .....TTGAACCAACTTTTGATTTGAGATGATGTTTGAAGGAAT 969
168 LysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGluY 184
    ||||| ||||| |||||
970 GAA..... 972
184 sValAlaValGlyAspValIleTyrIleGluAlaAsnSerGly..... 198
    ||||| ||||| |||||
973 .....GGGATTATGTCATGCTCCGCGAACAAACCTTCAGATGAC 1012
199 .....AlaValLysArgValGlyArgCysAspSer.....PheAla 210
    ||||| ||||| |||||
1013 TTGATCCTGCCCTTTGCGTCCAGAGCTTTGATGAATAAGTATGCTT 1062
211 ThrGluTyrAspLeuGluAlaGlnGluTyrValProIleProLysGly 227
    ||||| ||||| |||||
1063 GGCCTGCTGATGTTAAGGTCGTAAGCAATCTTG.....AA 1100
227 uValHis.....LysLysLysGluIleValGlnAspValThrLeuHisAsp 243
    ||||| ||||| |||||
1101 AGTTACGCTAAGAACACAGCCTTTACACAGATGTT.....GATT 1141
243 euAspAlaAlaAsnAlaGlnProGlnGly.....GlyGlnAspIleLeu 257

```



1142 TGAATTAGTGGGCTCAACAACCTCCAGCGCTTGTGTGTCGATATTAAAG 1191  
1143 ..::: 111 111 111:::  
258 SerLeuMeGlyGln.....MetLeuLysProArgLysThrGluLeu 272  
1192 AATGCTTGAATGAAACACACTTATGTTGCTGCTGCGC..... 1230  
272 rGluLysLeuArgGlnGluLeuLysValVal.....AsnArgTyr 287  
1231 ..... AATAAATCGAATATGATGCTTCAGATA 1258  
287 LeuAspGlnGlyLeuAglLeuValProGlyValLeuPheLeuAspGlu 303  
1259 TTGATTAACAGACAAGATATAGATTATTTGGTCGACCTTCAGAAAGATTAAG 1308  
304 ...ValHisMetLeuAspTleGluCysPheSerTyrLeuAsnArgAlaLe 319  
1309 ACAGTTTCCACAAAAMAAAGAAATTTGGTTGCTTAC...CATGAGGCGAGG 1355  
319 uGluSerProLeuSerProIleValIleLeuLanHisProGlyIleGlc 336  
1356 ACATTCACATTTGTTGT.....CTAGTCTTGTCGACTCTCGCGTTGGCC 1399  
336 yAsnAsnVal.....ArgGlyThrAspMetThrSerProHisGly 348  
1400 ATTAAGGTTTCAATTGTTACACACGCGCGCGAGGCGATACATGATGCA 1449  
349 IlePro.....ValAspLeuLeuAspArg 356  
1450 CTTCCTTAAGAGCATCAAAATGCTTCTATCTAAGAAAGATATGAAGAAGCA 1499  
356 gLeu.....ValIleIleArgT 362  
1500 ATTGGCTGGCTTAATGGGAGCAGTCGTGAGGAGAAATTTATCTTTAATG 1549  
362 hGluThrTyrGly.....ProThrGluMetIle 371  
1550 TCCAAATCATACAGACCTTCAAACGACTTTGAACAAGCCACCAAAATG... 1596  
372 GlnIleLeuValIleArgAlaGlnValGluGluLeuAspMetAspGlu 388  
1597 .....GCACGTGCATGGTTTACAGATGACGTATGAGTGAATAA 1634  
388 uSerLeuAlaTyrLeuGlnGlyGlnIleGlyGlnIleThrSerLeuArgHis 405  
1635 A.....CTTGGCCCAAGTA.....CAATATGAAGGAACAACATG 1666  
405 Le...IleGlnLeuIleSerProAlaSerValValSerLysThrAsnGly 420  
1667 CATCTCTTGTCGACAGAGTCTCTCAAAATCAATTTTCAGAAACAAACAGCT 1716  
421 ArgGluLysIleCysLysAlaAspLeuGlnGluValSerGlyLeuTyrLe 437  
1717 TATGAA.....ATTGAT...GAAAGGTTTGCTTATTATTATAA 1751  
437 uAspAlaLysSer...SerAlaArgLeuLeuGlnGlnGlnGluArgT 453  
1752 TGAGGCGACGAATAAAGCTGCTGAATATTATTCAGTCAAAATCGTGAACCTC 1801  
453 ytr 453  
:::  
1802 AC 1803

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-987-123-3

```
seq. documentation block:
: Sequence 3, Application US/08587123C
: Patent No. 6136557
:
: GENERAL INFORMATION:
: APPLICANT: Eli Lilly and Company
: TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE GENE SEQUENCE ftsH
: FILE REFERENCE: X-11762 Sequence Listing
: Patent No. 6136557
```

```

1 CURRENT APPLICATION NUMBER: US/08/9887,123G
2 CURRENT FILING DATE: 1997-12-08
3 PRIOR APPLICATION NUMBER: 60/036,281
4 PRIOR FILING DATE: 1996-12-13
5 NUMBER OF SEQ ID NOS: 4
6 SOFTWARE: Patentin Ver. 2.1
7 SEQ ID NO 3
8
9 LENGTH: 1959
10
11 TYPE: RNA
12
13 ORGANISM: Streptococcus pneumoniae
14
15 FEATURE
16
17 NAME/KEY: mRNA
18
19 LOCATION: (1)..(1959)
20
21 US-08-987-123-3

```

alignment\_scores:

Quality:	111.50	Length:	484
Ratio:	0.474	Gaps:	31
Percent Similarity:	48.554	Percent Identity:	23.140

```
alignment_block :
    rd 00 700 710
```

US-09-589-510-4 x US-08-987-123-3 .

Align seg 1/1 to: US-08-987-123-3 From: 1 to: 1959

[illegible]



```

2472 CATGGAATGTTCTGAGTAACGTTGCTTATCTGAGCTGTAAGTACT 2521
169 hValLlyGlyThrlsGlnleuysleuaspserlleTyr..... 177
2522 CCGCAAGATTAATGACCAATACGCTGATGATGCTCAATATTCGGCTTC 2571
178 .....AspAlaLeu1lelsglyleuysValAlaValGlyAspVa 190
2572 AGACCTCCCGGATGAGATATATGCGAAGAGACTAAG..... 2610
190 lIleTyr1leGluAlaAsnSerGlyAlaValLysArgValGlyArgCysa 207
2611 .....TACATGCGCGAAATGAGCGCTTAGAG..... 2637
207 spserPha1aThrlGluTyrAspLeuGluAlaGluGluTyrValPro1le 223
2638 .....CTACTGAGAGAGGTCTCCAGCAATATCTTAC.....ATA 2673
224 ProLysGlyGluValHisLysLysGluLeuValGlnAspValThr1e 240
2674 GCAGAGAGAGATATGAGAGAGACATATAACATCTCGAAGCTGCAGCAGC 2723
240 uHisAsp.....LeuAspAlaAlaAsnA 248
2724 TCTAGACAGAAGATCACCAGCAAAACGATATGATGATGAGTAGAG 2773
248 lAgInProGlnGlyGlyLysAsp.....lIleLeuSerLeuMetGlyGln 262
2774 CTAGACCTGAAGATATATAGAGATGATGCTTCTGCTTAAAGGCAAC 2823
263 MetMetLysProArgLysThrlleThrlGluLysLeuArgGlnlul 279
2824 TTCTTGAAGGCCAG.....GAAAAGCTTAGGAGAGACT 2858
279 easuLysValAlaAsnArgTyr1leAspGluGly1leAlaGluLeuValP 296
2859 TCTCAAG.....CAAGGACTTAGT..... 2877
296 roGlyValLeuPhe1leAspGluValHisMetLeuAsp1leGlyCysPhe 312
2878 ..GGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2916
313 SerTyrLeuAsnArgAlaLeuGluSerProLeuSerProLeuVal1lele 329
2917 .....ACCTGCGCATGAGAGAGCCAAAGAG.....GTTGCT 2951
329 uAlaThrAsnArgGly1leCysAsnValArgGlyThraspMetThrserP 346
2952 TGCTGATTAAGATAGAGAGATATTAACCTGACACTGCTGAA..... 2991
346 roHisGly1leProValAspLeuAspArgLeuVal1le1leArgThr 362
2991 ..... 2991
363 GluThrTyrGlyProThrGluMet1leGln1leLeuAla1leArgAlaG 379
2992 .....GGGCGTAATGAATATATTAATGCTTGAAGCAGCTTACGACA 3032
379 n 379
3033 G 3033
seq_name: /cgn2_6/ptcodata/2/lna/5B_COMB.seq:US-08-993-228-11
seq_documentation_block:
: Sequence 11, Application US/08993228
: Patent No. 5976838
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavallee, Edward R.
: APPLICANT: Racine, Lisa A.

```

```

APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,228
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-993-228-11

alignment_scores:
Quality: 108.50 Length: 542
Ratio: 0.448 Gaps: 25
Percent Similarity: 44.649 Percent Identity: 20.111

alignment_block:
US-09-589-510-4 x US-08-993-228-11 ..
Align seg 1/1 to: US-08-993-228-11 from: 1 to: 2837
23 GlyLeuGlyLeuAspAlaAsnGlyMetAla1leAlaLeuAla1leGlyPh 39
|||||.....
671 GGCATTGGCCGGACAGGCCGAGGATGGGGGACTTTCAGT..... 712
39 eValGlyGlnAlaAla..... 45
|||||.....
713 GTGCGAAGAACCACTGCAAGCTTAAAGATGAATGATGATGAAGT 761
46 ..ArgGluAlaAlaGlyLeuAla..... 52
|||||.....
762 TCAAGATGTGCTGCTGAGAGAGCCAGAGCTAGATCATGATGATTT 811
53 ValAspMet1leArgGlnLysLys.....MetAl 62
|||||.....
812 GTGATTTCTTGAACCAACCAAGCAGTATCAAGACTAGAGACATATAT 861
62 aglyArgAlaValLeuLeuAlaGlyProProAlaThrGlyLysThrAlaL 79
|||||.....
862 CCAAGAGGTGCGCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
79 eua1leuGly1leAlaGlnGluLeuGlySerLysValProPheCysPro 95
|||||.....
912 TAGCTAAGGCCACAGCCGAGAA.....GCCAATGTCCTTATCACC 955

```

```

96 MetValglySerGluValTyrSerSerGluValLysThrGluValLe 112
   :: |||||
956 GTTAGTGCATCTGAG.....TT 972
112 UmetGluAsnPharGargAlaIleGlyLeuArgIleLysGlu..... 126
   ::||| |||
973 TTTCGAGATGTCCTGTGTCGTGGCCCTGCTAGATCCGAGACTTATTTG 1022
127 .....AsnLysGlu.....ValTyrGluGlyValValThr 136
   ::|||::
1023 CCTTGCCTCGAAGAGATGCCCTTGCACTCCTTCATCGATGAATGAT 1072
137 GluLeuSerProGluGluValGluSerThrThrGlyLysValLys.. 152
   ::|||::
1073 GCGGTGGGAAGAGAGAGAGAGAGCACTTTCGAGGCAAGTGA 1122
153 .....Ser 154
1123 GGAGACACACTCAACCACTGCTGTGGAGATGATGTTTAAATACAA 1172
154 IeserHisValIleIleSerLeuLysThr..... 163
   ::|||::
1173 CAACAATGTCATTTTGGCCGACCAATGACAGACAGACAGACAGAC 1222
164 ValLysGly.....ThrLysGluLeuLysLe 172
   ::|||
1223 ATAAAGAGAGAGCTCTATTTTCAAGTTTCATCTCGACCCGCTAAACT 1272
172 uAspSerSerIleTyrAspAlaLeuIleLysGluLysValAlaVal... 187
   |||||::
1273 GGACAGTACCTCGAGAGAGATAAATTGGCAGAAACTGCACTCTTAA 1322
188 .....GlyAspValIleTyrIleGluAlaAsnSerGly 198
   ::|||
1323 CTCAGGCTTTTCAGTGTGATGTCCTATGCTGTATGAAGCTGCG 1372
199 AlaValLysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLe 215
   ::|||
1373 TTGATTCCTGCAAGCACTGTGTCAGATTCATTAATCAGAAACACTTGA 1422
215 uGluAlaGluGluTyrValProIleProLysGlyGluValHisLysLysL 232
   ::||| |||
1423 ACAGGCAATTGAGCGAGTG.....ATTGGGCTTAAAGAAAAA 1463
232 LysGluIleValGln.....AspValThrLeuHisAspLeu 243
   ::|||
1464 CGCAGGTTTCACACCTGAGAGAGAAAGACTGTGATACCCAGAGCA 1513
244 AspAlaAlaAsnAlaGln..... 249
1514 GGCCATGCGGTTGCCGCTGCTATCTGAGCAGCAGACCCGCTTTAA 1563
250 .....ProGlnGly..... 252
1564 GGTATGCATCATCCACGTGCAAGAGACTAGTTATGCTCAATTTTAC 1613
253 .....GlyGlnAspIleLeuSerLeuMetGly 261
1614 CAAAGAGACATACCTTATACCAAGAGCAGCTTTGGATAGAGATGTG 1663
262 GlnMetLeuLysProArgLysThrGlu.....IleThr 272
   ::||| |||
1664 ATGACTTTAGTGTGTCGATCTGAGAAATCTCTTTGAGAAATTTAC 1713
272 rGluLysLeuArgGlnGluIleAsnLysValValAsn.....A 285
1714 AACTGTGCTCAAGATGACTTGAGAAATCACTCAGAGCTATATGCC 1763
285 rGlyTyrIleAspGluGlyTyrIleAlaGluLeuValProGlyValLeuPheIle 301
1764 AAATGTTTCAGTTTGGCATGAATGAAGGTT...GGCAAAATCTCTCTT 1810
302 AspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnArgAl 318

```

```

1811 GACCTCCACGTCAGGGGAGATG.....GT 1836
   ||| ::|||
318 aLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArgGlyI 335
   ::|||
1837 ATTGGAGAAACCTTACAGTGA.....GCACTGCAAGATTTGA 1874
335 IecCysAsnValArgGlyThrAspMetThrSerProHisGlyTyrIleProVal 351
   |||
1875 TAGATCATGACATACCAATACTTATTATGATGCTTATTAAGAACAGTA 1924
352 AspLeuLeuAspArg.....LeuValIle 359
1925 GCTCTTCTCAGAGAAAGAAAGCTGACGTGAGAGGTTGCTTCTGTGTT 1974
359 eIleArgThrGluThrTyrGlyProThrGluMetIleGlnIleLeuAlaI 376
   ::|||
1975 GTTGAAGAAAGATTAATGATTAAGATGATVAGTTGAACCTTTGGGCC 2024
376 IeArgAlaGlnValGluGluIleAspMetAspGluGluSerLeuAlaTyr 392
   |||
2025 CCAGACCATTTGGGAAAAA...TCTACTATGAAGATTTGTGAAAGC 2071
393 LeuGlyGluIleGlyGlnGlnThrSerLeuArgHisAlaIleGlnLeuI 409
   |||
2072 ACTGGCAGCTTGATGAGAGACACTCAGCT..... 2101
409 eSerProAlaSerValAlaSerLysThrAsnGlyArgGluLysIleCysL 426
   |||
2102 ...CCAGAGGCTTAAAGACTGGAACAGAGAGCGGAAAAAG..... 2140
426 LysAlaAspLeuGluGluValLysSerGly 434
2141 ....GAGAAAGAGAGACCCCGGGGT 2161

```

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-103-840A-2

seq\_documentation\_block:

```

: Sequence 2, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTNER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4403765
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

alignment\_scores:

```

Quality: 133.50 Length: 351
Ratio: 0.746 Gaps: 18
Percent similarity: 50.997 Percent identity: 22.222

```

alignment\_block:

US-09-589-510-4 x US-09-103-840A-2 ..

Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765

```
36 AlaIaagIlyPheValIGlyGlnAlaAlaIaArgIuValaIaagIlyLeuAl 52
||| |||.....:~::~||| |||
1669511GCCGAGGGGCTGGCCGCGAGGTGCACACGCTGAGCGGCATCTTCCA 1669560
52 aVal.....AspMetIleArgGlnIlyLysMetA 62
||| |||.....:~::~||| |||
1669561GGTCACAGCGGATATGCTGCGGCCAGAGCAGCTGTGTGAGCGGATGCTCG 1669610
62 IaGIyArGAla.....ValLeuLeuAlaGIyProProAlaThr 74
||| |||.....:~::~||| |||
1669611TCGGCCCTGCTGCACAGGCGCATGTGCTGTGAGGCGCTTCCCGGCGTG 1669660
75 GLyIyThrAlaLeuAlaLeuGIyIleAlaGlnGluLeuGIy..... 88
||| |||.....:~::~||| |||
1669661GGCAGAGCTTGGCGGTGAGACCTTGTGCTGGGTGCTGCGCGGACATT 1669710
89 .SerIyValIProPheCysProMetValIGlySerGIuValIySerSerG 105
||| |||.....:~::~||| |||
1669711TTCGGCATCCAGTTCAACCCG.....G 1669733
105 LuValIyLysIyThrGIuValIleuMetGIuAsnPheArGAlaIleGIy 121
||| |||.....:~::~||| |||
1669734ATCTGTGGCCACCGCATC.....ATCGGG 1669759
122 LeuArGAla.....LysGIuAsnLysGIuValIyGIyGIuValIThrGI 137
||| |||.....:~::~||| |||
1669760ACGCGCATCTACCGCAGGAGGAGCAATTCAC.....ACCGA 1669800
137 uLeuSerProGlnGIuAlaGluSerThrThrGIyGIyAlaLysSerI 154
||| |||.....:~::~||| |||
1669801ACTCGAGACCGGTGTGGCAACTCTCG.....CTCGCGACAGAGA 1669841
154 LeSerHisValIleIleSerLeuLysThrValIyLysGIyThrLysGIuLeu 170
||| |||.....:~::~||| |||
1669842TTCACCCGGCT.....CCGGCG 1669858
171 LysLeuAspSerSerIleTyAspAlaLeuIleLysGIuLysValAlaVa 187
||| |||.....:~::~||| |||
1669859AAGTGGCAGTGGCGTCTGAGATGACGAGAGGCCCATGTGTCAT 1669908
187 IGIyAspValIleTyIleGluAlaAsnSerGIyAlaValIyArGAlG 204
||| |||.....:~::~||| |||
1669909CGGGATAGACCTTCCGATGCCACCGCTCTGTGTG..... 1669948
204 IyArGcYAspSerPheAlaThrGIuTyIAspLeuGIuAlaGIuGIuTy 220
||| |||.....:~::~||| |||
1669949.....ATGGCGACGCAACCCGATCGACGACGAGCGGCTC 1669984
221 ValProIleProLysGIyGIuValHisLysLysGIuLeuValGlnAs 237
||| |||.....:~::~||| |||
1669985TACCCGCTACCGGAGCGCAA.....CGGGA 1670010
237 pValIThrLeuHisAspLeuAspAlaAlaAsnAlaGlnProGlnGIyGI 254
||| |||.....:~::~||| |||
1670011CCGCTTCTTCAAGATCAACGCGCTACCCGCTACCGCGAAGAGAGC 1670060
254 InaSPIleLeuSerLeuMetGIyLysMetLysProAlaGIyLysThrGI 270
||| |||.....:~::~||| |||
1670061CGCAATCATCTACCGTATGGT...GTACCCCGCGCAGGCGCAAGACG 1670107
271 IleThr.....GluLysLeuArG...GlnGIuLysLysValVa 283
||| |||.....:~::~||| |||
1670108ATCTCGAGACAGCGGCGACCTGCTGCGCTGAGAGATA.....GGGCG 1670151
283 InaArGIyTrIleAspGIuLysIleAlaGIuLeuValProGIyValLeuP 300
||| |||.....:~::~||| |||
1670152CAACAACCTTCTCACACACGCGGTGTGACTATGTGCTGAGTCTCT 1670201
300 heIle.....AspGIuValHisMetLeuAspIleGIuLysPhe 312
||| |||.....:~::~||| |||
1670202TTCGCCACCCGCAACCCGAGAGTGGGATGAAGAACGAGCTGG 1670251
313 SerTyIleuAsnArGAlaLeuGIuSerProLeuSerProIleValIlele 329
```

```
1670252GTTCGCTTGGC.....GCATCCCCGCGTGCCTGCTGGCATCAT 1670292
329 uAlaThrAsnArGIyTrIleCysAsnValArGIyThrAspMetThrSerP 346
||| |||.....:~::~||| |||
1670293CGCCCGCGCAGTCCCTGCGCTGTGCGGCGGTGACTATGTTCATCC 1670342
346 rOHIsGIy.....IleProValAspLeuLeuAspArgLeuVal 358
||| |||.....:~::~||| |||
1670343CGCAACACTCATTCGAGGTATTCCTGATGTCTGTGAGACACCGGCTCGTG 1670392
359 Ile 359
1670393CTC 1670395
seq_name: /cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-074-579-2
seq_documentation_block:
: Sequence 2, Application US/09074579
: Patent No. 6001596
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Guejler, Karl J.
: APPLICANT: Patterson, Chandra
: TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
: TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/074,579
: FILING DATE: HEREMITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Ceirone, Michael C
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PF-0505 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: UTRNOT02
: CLONE: 688183
: US-09-074-579-2
alignment_scores:
Quality: 104.50 Length: 484
Ratio: 0.525 Gaps: 23
Percent Similarity: 41.116 Percent Identity: 17.975
alignment_block:
US-09-589-510-4 x US-09-074-579-2 ..
```

Align seg 1/1 to: US-09-074-579-2 from: 1 to: 3636

```

93 PheCysProMetValIglySerGluValTyrSerSerGluValLysThr 109
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
380 TTCACCTATGCTTATTGGAGACAGAGTGATATCAGGCCCAATTAACAGAG 429
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
109 rGluValLeuMetGluAsnPheArgAlaIleGlyLeuArgIleLysG 126
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
430 AGAA.....AAGAGAGTGATAGGGTAAAG 458
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 LuAsnLys.....GluVal 130
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 AGAAAGAGATAAACCACAGAGAAATGAGAGAAAGGCGACTGAATA 508
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
131 TyrGluGlyGluValThrGluLeuSerProGluGluAla..... 143
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
509 TTCAGAGCTTCTGCAGTATCCACAGACAGACAAACCCCTTTTTCCT 558
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 .....GluSerThrThrGlyGlyTyrAlaLysSerI 154
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
559 GAGTTATGAGAGGCTCTGCAGAGGCGCTGGCAAGTACGACACAGC 607
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 IeSerHisValIleIeSerLeuLysThrValLysGlyThrLysGluLeu 170
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 .....ATCAGCGTGCGGCCCCACAGAGCTGCGGAGGCGTGC 643
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
171 LysLeuAspSerSerIleTyrAspAla..... 179
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
644 AGCGTGACCGTAATATCTCGAGAGCGCGGCATCGCATCCCTGGAGGT 693
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 ..... 179
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
694 GCTGCCCTTCACACAGACAGAGGCGAGTGGCGCGGGAGAGATG 743
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 .....LeuIleLysGluLysValAlaIleValGly 188
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
744 ATTCTGGGCTCCCTCTACTGTATTAACCAAAATGAACATTGTGCC 793
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 AspValIleTyr.....IleGluAlaAsnSerGlyAlaValLysArg 203
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
794 AACATATTTTAACTACTAGTACACAAACAGCAGATTCCTCCAGAA 843
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 IglyArgCysAspSerPheAlaThrGluTyrAspLeuGluAlaGluLys 220
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
844 TGAATTTTGGAGACTTATCATATGATATGCTCAATAGAGACAGA 893
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 yVal..... 221
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
894 GCATTGGGACATCCAGGTTCTAATAGCTATTGTGCACTACTTGTCT 943
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 .....ProIleProLysGly..... 226
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
944 CCTAAGACCTTCCTCTTACCCAGAAATGTGATTTGCTGTGACAG 993
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
227 .....GluValHisLysLysLysGluIleValG 236
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
994 CAGTCTCTATGATGGAGAACCAACTCCGGCAGACCAAGATGCCCTCT 1043
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 IAspValThrLeuHisAspLeuAspAlaAlaAsnIleProGluGly 252
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1044 TCACAAT...CTCCATGACCTC.....CGACCC..... 1069
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 GlyIleAspIleLeuSerLeuMetGly..... 261
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1070 ...CAGGACCGTTTCAGTATCATTTGATTTTCCAAACCGGATCAAGTATG 1116
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 ..... 261
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1117 GAAGACCACTTGATATCATCTCCACAGACATCAGGAGTGGGAAG 1166
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 .....GlnMetMetLysProAlaGlyThrGluIleThrGluLys 274
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1167 TGTACATTCACATATATGTCACCCACTGAGGACACAGACATCAAGGGGCC 1216
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
275 LeuArgGlnGluIleAsnLysValAlaAsnArgTyrIle...AspGluG 290
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1217 CTCACAGAGGCCATC...AGGCTCCCAACAGATGAGTGCCCAACATGG 1263
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 yIleAlaGlu.....LeuValProGlyValL 299
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1264 CATTTGAGACCGAGCGGTGCTCCATCGTTCTTCCAGAGATGGGAAGC 1313
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 eupHeIleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeu 315
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1314 CCAGGTCGGGAGACGACACCTCAAGATC.....CTC 1348
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 AsnArgAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAs 332
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1349 AACACACCCCGAGAGGCCGCCGAGGACCAAGTCTCATCTTACCATTTGG 1398
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 nArgGlyIleCysAsnValArgGlyThrAspMetThrSerProHisGlyI 349
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1399 CATGGC.....AAGCAGC 1412
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 IeProValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyr 365
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1413 TGGACTTCAGGCTGCTGGAGAACTGTGCTG.....GAGAACTGT 1453
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 GlyProThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluG 382
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1454 GGCCTCACACGCGCGGTGCAC.....GAGGA 1479
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 uIleAspMetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGln 399
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1480 GGAGAGACGAGCTCGCAGCTCATCGGCTTCTAGCATGAATC.....A 1523
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 IuThrSerLeuArgHisAlaIleGlnLeu...IleSerProAlaSerVal 414
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1524 GGACCCCGCTCCTCTGTGACATCGCATCGCATATATCCCGACGCTCAGT 1573
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 ValSerLysThr.....AsnGlyArgGly 423
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1574 GTGAGGCCACACAGACCTGTGCCCACTACTTCAACGCGCTGGAGAT 1623
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
423 sIleCysLysAlaAspLeuGluGluValSerGlyLeuTyrLeuAspAla 440
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1624 CATC.....ATTGCGGGAGAGCTGGTGACACAGA 1652
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
440 ys 440
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1653 AG 1654

```

seq\_name: /cgm2\_6/ptodata/2/ina/6b\_COMB.seq:US-09-388-774-2

seq\_documentation\_block:

Sequence 2, Application US/09388774

Patent No. 6228991

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Guegler, Karl J.

APPLICANT: Patterson, Chandra

TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE

TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

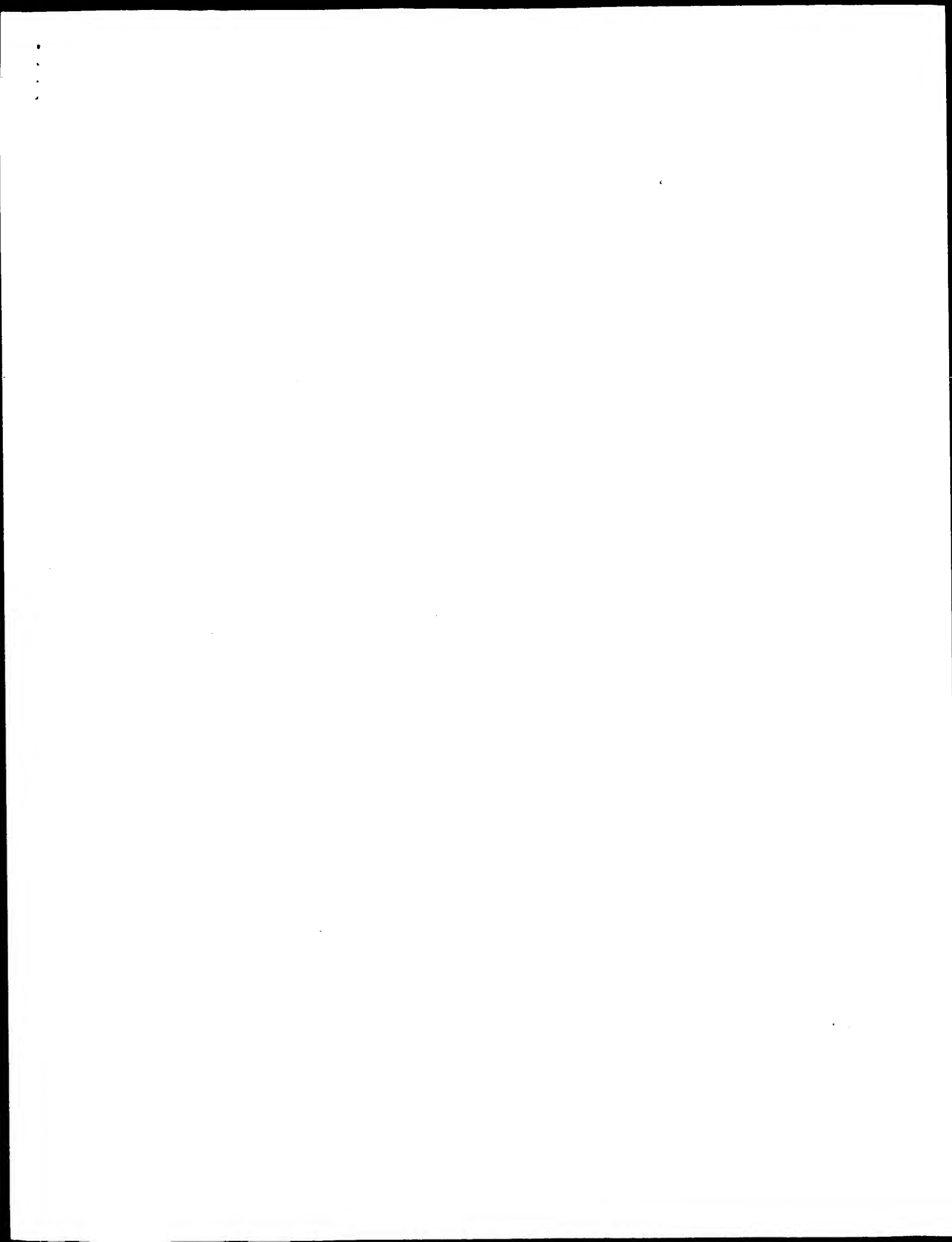


16 A1aThrHisThrlleLysGlyLeuGlyLeuAspAlaAsnGlyMetAl 32

[illegible]



984 GAAAGAAAACAAATGTACAGAGAGAGGCTGAACCTGATGGAAACAGTGG 1033  
290 YLEAlaGluLeuValPro 296  
|::| | | | | | | | | |  
1034 TCTGGCTGAGTTGACCCA 1052



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2001, 06:56:39 ; Search time 25.58 seconds

(without alignments)  
2353.354 Million cell updates/sec

Title: US-09-589-510-4

Perfect score: 2263  
Sequence: 1 MREIEVQSTKKQRIATHH.....YLDKSSARILQEQEERYIT 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SPREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.unclassified:*
14: sp.virus:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2039	90.1	458	10 Q9M326	Q9m326 cicer ariet
2	1994	88.1	458	10 Q9PMB9	Q9Pmb9 arabidopsis
3	1723.5	76.2	456	4 Q9Y265	Q9Y265 homo sapien
4	1722.5	76.1	456	11 Q35753	Q35753 rattus norv
5	1719.5	76.0	456	13 Q9DE26	Q9de26 xenopus lae
6	1655.5	73.2	456	5 Q9VH07	Q9vh07 drosophila
7	1643.5	72.6	456	5 Q9NH53	Q9nh53 drosophila
8	1517.5	67.1	463	3 Q03940	Q03940 saccharomyc
9	1311	57.9	458	5 Q17607	Q17607 caenorhabdi
10	989.5	43.7	465	3 Q94692	Q94692 schizosacch
11	989	43.7	462	13 Q9DE27	Q9de27 xenopus lae
12	989	43.7	481	5 Q9V3K3	Q9v3k3 drosophila
13	986.5	43.6	441	1 Q9V1R4	Q9v1r4 pyrococcus
14	975	43.1	441	1 Q59468	Q59468 pyrococcus
15	974	43.0	463	11 Q9WTM5	Q9wtm5 mus musculu
16	973	43.0	463	4 Q9Y230	Q9Y230 homo sapien
17	963.5	42.6	469	10 Q9FCW0	Q9fCW0 arabidopsis
18	942.5	41.6	449	1 Q28462	Q28462 archaeoglob
19	932.5	41.2	471	3 Q12464	Q12464 saccharomyc

20	928	41.0	473	10 Q9M2X5	Q9m2x5 arabidopsis
21	895	39.5	442	4 Q9Y361	Q9Y361 homo sapien
22	888.5	39.3	473	1 Q9YF85	Q9Yf85 aetopyrum p
23	863	38.1	448	5 Q9G2H2	Q9g2h2 caenorhabdi
24	417	18.4	118	5 Q9GK8	Q9gk8 leishmania
25	299.5	13.2	112	1 P96020	P96020 sulfolobus
26	180.5	8.0	795	1 Q9V1N2	Q9v1n2 pyrococcus
27	179.5	7.9	798	1 Q59515	Q59515 pyrococcus
28	177	7.8	840	1 Q58419	Q58419 pyrococcus
29	171	7.6	835	1 Q9Y2X2	Q9Y2x2 pyrococcus
30	168	7.4	840	1 Q9Y2Z7	Q9Y2z7 pyrococcus
31	166.5	7.4	733	1 Q28972	Q28972 archaeoglob
32	159	7.0	1251	10 Q9MA34	Q9ma34 arabidopsis
33	153.5	6.8	699	1 Q9Y910	Q9Y910 aetopyrum p
34	148.5	6.5	726	1 Q9YC86	Q9Yc86 aetopyrum p
35	146.5	6.5	811	1 Q28182	Q28182 archaeoglob
36	142.5	6.3	614	8 Q19922	Q19922 cyanidium c
37	141.5	6.3	892	11 Q9Z2K7	Q9Z2k7 mus musculu
38	141	6.2	620	4 Q9H8W7	Q9H8w7 homo sapien
39	140	6.2	737	1 Q9HHW2	Q9hhw2 halobacteri
40	139.5	6.2	732	1 Q27676	Q27676 methanobact
41	138	6.1	809	3 Q60058	Q60058 schizosacch
42	137.5	6.1	709	2 Q9JTS1	Q9Jts1 neisseria m
43	136.5	6.0	943	5 P91638	P91638 drosophila
44	136.5	6.0	944	5 Q9VS62	Q9VS62 drosophila
45	136	6.0	759	1 Q9HPU1	Q9hpu1 halobacteri

## ALIGNMENTS

## RESULT 1

Q9M326 PRELIMINARY; PRT; 458 AA.

```

AC Q9M326;
ID Q9M326;
DI 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE PUTATIVE RUV DNA-HELICASE.
OS Cicer arietinum (chickpea). (Garbanzo).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Cicer.
OX NCBI_TaxID=3827;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
RA Dopico B., Esteban R., Labrador E.;
RT "A RuvB DNA-helicase like protein is expressed in chickpea
epicotyls."
RT Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RL EMBL; AJ276264; CAB76908.1; -.
DR InterPro: IPR001984; -.
DR InterPro: IPR003593; -.
DR PRINTS: PR00830; ENDOLAPTASE.
DR SMART; SM00382; AAA; 1.
KW Helicase.
SQ
SEQUENCE 458 AA; 50262 MW; 61BAD73060376764 CRC64;

```

Query Match 90.1%; Score 2039; DB 10; Length 458;

Best Local Similarity 87.5%; Pred. No. 1.1e-109; Matches 398; Conservative 33; Mismatches 24; Indels 0; Gaps 0;

QY	1	MREIEVQSTKKQRIATHHINKIGLDANGMAIALAGFVGQAAREAGLAVMIROKK 60
DB	4	MKEIEVQSTKKQRIATHHINKIGLEVGSKALPFASGFVGQAAREAGLAVDMIRKK 63
QY	61	MAGRNVLLAPPAATGKTALALAGIAGELGSKVPPCPVAVGSEVYSSVKKTEVLMENFRRI 120
DB	64	MACKALLLWGPPTGKTALALAGIAGELGSKVPPCPVAVGSEVYSTEKKTEVLMENFRRI 123
QY	121	GLIRKKNKEVEGEVTELSPEAESEFTTGGYAKSISHTVITSLTKVGTOKLIDDSIYDAL 180

Db	124	GLRIKENKEVEGEVTELTSPETESLINOGYKSTSHVILGKTYKGRKOLKLDPTIYDAL	183
Qy	181	IKEKVAVDVITYEANSQAVRVRGRCSFATEYDLAEAEVVPJPKGEVHKKEIVDVTL	240
Db	184	IKKEVAVDVITYEANSQAVRVRGRSDAFATEPFLDEAEVVPJPKGEVHKKEIVDVTL	243
Qy	241	HDLDAAANOPOGGODIILSLMOQMKPKRTEETELKROEINKVVRNYIDEBIAELVRYLFL	300
Db	244	HDLDAAANRPOGGODIILSLMOQMKPKRTEETDRLROEINKVVRNYIDEGVAELVRYLFL	303
Qy	301	IDEVAMLDIECFSYLNRALESPLSPYILATNRGICVNRGCTDMTSPHGIIVDLDRLVIT	360
Db	304	IDEVAMLDIMECFSTYLNRALESPLSPYIFANNRGICVNRGCTDMTSPHGIIVDLDRLVIT	363
Qy	361	RTEYTGPELMTIOILAIRAQVEEIDMDEBSLAVLGEIGQOTSLRAHAIOLISPASVSKTNG	420
Db	364	RTOFGPAPETIIIOILAIRAQVEELVAVDEBSLAFGEIGQRTSLRAHAIOLISPASVAKING	423
Qy	421	REKICKADLEEVSGLYIDAKSARLLEOQOERYIT	455
Db	424	RDNICKADLEETCSLYIDAKSARLLEOQOERYIT	458

	RESULT	2	
09FMR9			
ID	Q9FMR9	PRELIMINARY;	PTN; 458 AA.
AC	Q9FMR9		
DT	01-MAR-2001	(TREMBLrel. 16, Created)	
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)	
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)	
DE	ROY DNA-HELICASE-LIKE PROTEIN.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-COLUMBIA;		
RX	MEDLINE-98162728; PubMed-9501997;		
RA	Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,		
RA	Tabata S.;		
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. III.		
RT	Sequence features of the regions of 1,191,918 bp covered by seventeen		
RT	physically assigned pl clones."		
RL	DNA Res. 4:401-414(1997).		
DR	EMBL; AB007651; BAB08331.1; -.		
DR	Helicase.		
GO	SEQUENCE	458 AA;	50323 MW; 7DA33A99AFA365F CRC64;

Query Match	88.1%;	Score 1994;	DB 10;	Length 458;
Best Local Similarity	85.1%;	Pred. No. 4.4e-107;		
Matches 387;	Conservative 39;	Mismatches 29;	Indels 0;	Gaps 0;

OY	1	MRIEVOSTKOKRIATHHTHKGIGDANAMATLALAFGQGAARPAALAYADMIROKK	60
Dd	4	VRIEIOSTAKKORATHHTHKGIGLEPTPIPLKANGFGVQLFARRPAGLVYDMIRÖKK	63
OY	61	MAGRALLAAPPATGKTALALGIAOELGSKVFPFCPMWSEVYSSEVKTEVLMEENERRAI	120
Dd	64	MAGRALLALBPBGTKTALALGISOELGSKVPCPMWSEVYSSEVKTEVLMEENERRAI	123
OY	121	GLRIRKENVEGEVTETLSPEEAESTTGGAKSISHVISLTAKTVKGTOKJOLDSITYDAL	180
Dd	124	GLRIRKEKEVEGVTELSPETESITLGVGSGKSISSHVITILKTIVKGTIKHLKLDPITYDAL	183
OY	181	IKEKVAAGDVIIYIANSNGAVRWGRCDSPAFTEDYLEAEZYVPIPKGEVHKKELEYODVYL	240
Dd	184	IKEKVAAGDVIIYIANSNGAVRWGRSDAPAFTEPDLEAEZVYLPFKGBGVHKKEKIYODVYL	243
OY	241	HDLDAANAQPOGGODIISLMCGOMMKPRKTTETTERKLOELINKVANRYIIDBSIAELVPGLFE	300

Db	244	QDLDAHNRPOGGODIISLMQOMKPKRTETITDLROEINKVNRVYIDEGVAELVPEVLF	303
QY	301	IDEVMLDIECFSYLNRALESPLSPVLTATNRGICVNRGDMSPGAPIDLLDRLVII	360
Db	304	IDEVMLDIECFSYLNRALESLSLSPVLTATNRGICVNRGDMSPGAPVITDLDRLVIT	363
QY	361	RTEYGTPEMIIQILAIRQVEEIMDEBSLAYLGEIGQOTSLRHAJOLISPAVSXKTG	420
Db	364	RQIQTDPSEMIQIILAIRQVEELVYDEECVLVLGEIQGRISLRHAVALDSPASTVAKMG	423
QY	421	REKICKADLEEVSGLYDPAKSSARLLQEOOERYIT	455
Db	424	RDNICKADIEEYTSIYDPAKSSAALLHQEQPKYIS	458

ID	Q9Y265	PRELIMINARY:	PRT:	456 AA.
AC	Q9Y265;			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)			
DE	ERYTHROCYTE CYTOSOLIC PROTEIN OF 54 KDA, ECP-54.			
CN	Trp49.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BONE MARROW.			
RX	MEDLINE=98447618; PubMed=9774387;			
RA	Kudlick M., Salzer M., Prohaska R.;			
RT	"An eukaryotic RuvB-like protein (RUVBL1) essential for growth.";			
RL	submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98447618; PubMed=9774387;			
RA	Qiu X.B., Lin Y.L., Thome K.C., Plan P., Schlegel B.P., Weremowicz S.S.,			
RA	Parvin J.D., Dutta A.;			
RT	"An eukaryotic RuvB-like protein (RUVBL1) essential for growth.";			
RL	J. Biol. Chem. 273:27786-27793(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=COLORECTAL ADENOCARCINOMA;			
RX	MEDLINE=99061945; PubMed=9843967;			
RA	Bauer A., Huber O., Kemler R.;			
RT	"Pontin5, an interaction partner of beta-catenin, binds to the TATA			
RL	box binding protein.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 95:14787-14792(1998).			
RP	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER.			
RX	MEDLINE=98249797; PubMed=9588198;			
RA	Makino Y., Mimori T., Kolke C., Kanemaki M., Kurokawa Y., Inoue S.,			
RA	Kishimoto T., Tamura T.;			
RT	"Trp49, homologous to the bacterial DNA helicase RuvB, acts as an			
RL	autoantigen in human.";			
RN	Biochem. Biophys. Res. Commun. 245:819-823(1998).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RX	MEDLINE=99032802; PubMed=9813143;			
RA	Holzmann K., Gerner C., Korosec T., Poeltl A., Grimm R., Sauermann G.;			
RT	"Identification and characterization of the ubiquitously occurring			
RL	nuclear matrix protein NMP 238.";			
DR	Biochem. Biophys. Res. Commun. 252:39-45(1998).			
DR	EMBL: Y18418; CAB46271.1; -			
DR	EMBL: AF070735; AAC77819.1; -			
DR	EMBL: AF099084; AAD04427.1; -			
DR	EMBL: AB012122; BAA28169.1; -			
DR	EMBL: AJ010058; CAA08986.1; -			





09NH53  
ID 09NH53 PRELIMINARY; PRT; 456 AA.  
AC 09NH53;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE PONTIN.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Phyllozoa; Drosophilidae; Drosophila.  
ON NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chauvet S., Ueseglio F., Missod R., Aragnol D., Pradel J.;  
RT "Control of MtC signalling by the RuvB-like Pontin and Reptin  
RT proteins.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF233278; AAF4341.1; -  
DR FlyBase; FBgn0040078; pontin.  
DR InterPro; IPR003593; -  
DR SMART; SM00382; AAA; 1.  
DR SEQUENCE 456 AA; 50215 MW; D2FA09186A1E9B0E CRC64;

Query Match 72.6%; Score 1643.5; DB 5; Length 456;  
Best Local Similarity 68.8%; Pred. No. 6.3e-87;  
Matches 313; Conservative 74; Mismatches 67; Indels 1; Gaps 1;

QY 1 MRIEVOSTSKORATHTHKGIGDANGMAIALAGFVGQAARARAGIAVDMIRK 60  
DB 1 MKIEVNSTVTORAAHSHVKGGLDEVGAAVSHAGLQKARERAGIIVLISKK 60  
QY 61 MAGRAVLLAGPATGKTALALGIAOELSGKVPFCPMVGSSEYSEVKTTEVLMENFRRAI 120  
DB 61 MAGRAVLLAGPATGKTALALGIAOELSGKVPFCPMVGSSEYSEVKTTEVLMENFRRAI 120  
QY 121 GLRIKENEVEYEGVETELSPERASTGGYAKSISHVYISLKYTKYQKLDSDSYDAL 180  
DB 121 GLRIKENEVEYEGVETELSPERASTGGYAKSISHVYISLKYTKYQKLDSDSYDAL 180  
QY 121 GLRIKENEVEYEGVETELSPERASTGGYAKSISHVYISLKYTKYQKLDSDSYDAL 180  
DB 121 GLRIKENEVEYEGVETELSPERASTGGYAKSISHVYISLKYTKYQKLDSDSYDAL 180  
QY 181 IREKAVGDVITYEANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 240  
DB 181 IREKAVGDVITYEANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 240  
QY 181 IREKAVGDVITYEANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 240  
DB 181 IREKAVGDVITYEANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 240  
QY 241 HDLDAANAPOGGODILSLMGOMKPKRTETETKLOEINKVNRVYIDEGIAELVPGVLF 300  
DB 241 HDLDAANAPOGGODILSLMGOMKPKRTETETKLOEINKVNRVYIDEGIAELVPGVLF 300  
QY 301 IDEVHMDIECTSYNRALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 359  
DB 301 IDEVHMDIECTSYNRALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 359  
QY 301 IDEVHMDIECTSYNRALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 359  
DB 301 IDEVHMDIECTSYNRALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 359  
QY 360 IRTETGPTETIOIATIRAOVEIIMDEESLAYIGEIGQOTSLRHAIOISPASVYKTN 419  
DB 360 IRTETGPTETIOIATIRAOVEIIMDEESLAYIGEIGQOTSLRHAIOISPASVYKTN 419  
QY 361 IRTLLYSTADMGQITKRAQTEGLOLEENAFTRLSIEIGSSFLRYAVOULPFAHGMCKV 420  
DB 361 IRTLLYSTADMGQITKRAQTEGLOLEENAFTRLSIEIGSSFLRYAVOULPFAHGMCKV 420  
QY 420 GREKICKADLEEVSGIYLDKASSARLLOEQERYT 454  
DB 420 GREKICKADLEEVSGIYLDKASSARLLOEQERYT 454  
QY 421 GRMOISKDIEDVHSLFLDAKRSKHLSEKNKFM 455  
DB 421 GRMOISKDIEDVHSLFLDAKRSKHLSEKNKFM 455

RESULT 8  
ID 003940 PRELIMINARY; PRT; 463 AA.  
AC 003940;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE HYPOTHETICAL 50.5 KDA PROTEIN.  
GN RVB1.  
OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
ON NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB972;  
RA Oliver K., Harris D.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB972;  
RA Barrell B., Rajandream M.A., Walsh S.V.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z48784; CAA8704.1; -  
DR SGD; S0002598; RVB1.  
DR InterPro; IPR003593; -  
DR SMART; SM00382; AAA; 1.  
DR Hypothetical protein.  
SO SEQUENCE 463 AA; 50453 MW; 93D6ECH7006AF0F2 CRC64;

Query Match 67.1%; Score 1517.5; DB 3; Length 463;  
Best Local Similarity 65.7%; Pred. No. 1.1e-79;  
Matches 286; Conservative 81; Mismatches 67; Indels 1; Gaps 1;

QY 14 RITHTHIGKLGIDANGMAIALAGFVGQAARARAGIAVDMIRK 73  
DB 23 RITHTHIGKLGIDANGMAIALAGFVGQAARARAGIAVDMIRK 73  
QY 74 TGTALALGIAOELSGKVPFCPMVGSSEYSEVKTTEVLMENFRRAI 133  
DB 83 TGTALALGIAOELSGKVPFCPMVGSSEYSEVKTTEVLMENFRRAI 133  
QY 134 EVELTEPEDEENPLGQYKTSISHVYISLKYTKYQKLDSDSYDAL 193  
DB 143 EVELTEPEDEENPLGQYKTSISHVYISLKYTKYQKLDSDSYDAL 193  
QY 194 EANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 253  
DB 203 EANSAGVRYKRGDSFATEYDEAEYVPIPKGEVHKKEIYODVTL 253  
QY 254 QDILSLMGOMKPKRTETETKLOEINKVNRVYIDEGIAELVPGVLF 313  
DB 263 QDILSLMGOMKPKRTETETKLOEINKVNRVYIDEGIAELVPGVLF 313  
QY 314 YLNALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 372  
DB 323 YLNALESPLPIVILATNRGICNVRG-TDMTSPHGIPVDLDRIVT 372  
QY 373 ILAIRAOVEIIMDEESLAYIGEIGQOTSLRHAIOISPASVYKTN 432  
DB 383 ILAIRAOVEIIMDEESLAYIGEIGQOTSLRHAIOISPASVYKTN 432  
QY 433 SGLYLDKASSARLLO 447  
DB 443 KLFLLDKASTKILE 457

RESULT 9  
ID 017607 PRELIMINARY; PRT; 458 AA.  
AC 017607;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE C27H6.2 PROTEIN.  
GN Caenorhabditis elegans.  
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;  
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.  
ON NCBI\_TaxID=6239;  
RN [1]  
SO SEQUENCE FROM N.A.

RA Gardner A.:  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed-7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Zeller M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner M., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden K., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
 RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: Z81042; CAB02793.1; -  
 DR InterPro: IPR003593; -  
 DR SMART: SM00382; AAA; 1.  
 SQ SEQUENCE 458 AA; 50616 MW; F801E34FADCC0EE2 CRC64;

Query Match 57.9%; Score 1311; DB 5; Length 458;  
 Best Local Similarity 57.8%; Pred. No. 8.3e-68;  
 Matches 262; Conservative 78; Mismatches 91; Indels 22; Gaps 5;

OY 3 IEEVOSTSK-ORATHTHRIKGLIDANGM-AIALAGFVGQAAREAGLAVDMIRQK 60  
 DB 20 IEEVPTKQKRIASHVHVGIDTETQAHNEAGFVGQAARFRASTIVMIRKLC 79  
 OY 61 MAGRAVLGAPGATKTAALGIAOELGSKVPCPMGSEVYSEVKKTEVLMENFRRAI 120  
 DB 80 MAGRAVLGAPGATKTAALGIAOELGSKVPCPMGSEVYSEVKKTEVLMENFRRAI 139  
 OY 121 GLRIKENKEVEGETEYTESPEAESTTGGYAKSISHVTSIKTKYKQKOLDSITDAL 180  
 DB 140 GLRIKENKEVEGETEYTESPEAESTTGGYAKSISHVTSIKTKYKQKOLDSITDAL 198  
 OY 181 IEEVAVGVYIYIANSAGVAVKRCDFATEYDLEAEVYPIPKGVHKKELVODVTL 240  
 DB 199 LKQHEVEDVYIYIANSAGVAVKRCDFATEYDLEAEVYPIPKGVHKKELVODVTL 258  
 OY 241 HDLDAANAPGCGO-DILSLMGOMKPRKTEITEKLOEINKVYNNYIDEGIAELVPGVL 299  
 DB 259 HDDDIANARPGCGO-DILSLMGOMKPRKTEITEKLOEINKVYNNYIDEGIAELVPGVL 318  
 OY 300 FIDEVHMLDIECSYLRALSPILVILATNGICNVGTDMTSPHGPVLDRLRYI 359  
 DB 319 FIDEVHMLDIECSYLRALSPILVILATNGICNVGTDMTSPHGPVLDRLRYI 378  
 OY 360 IRFTYGTPTMIQILAIRAOVEELIDMEESLAYLGIGQOTSLRHAIIOLSPASVSKTN 419  
 DB 379 IPTKKNIEEDIRKILVHRTAEVNOFEKAFDLTRLCQTC----- 420  
 OY 420 GREKICKADLEEVSGYLDKSSARLLQEOER 452  
 DB 421 GREVIEVEDVDRCTKLPMNGESLKAABEMRO 453

RESULT 10  
 OY 094692 PRELIMINARY: PRT; 465 AA.  
 AC 094692;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 GN SPB83.08.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA Wood V., Rajandream M.A., Barrell B.G., Bothe G., Pohl T.,  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE AAA FAMILY OF ATPASES.  
 DR EMBL: AL035536; CAB36870.1; -  
 DR InterPro: IPR003593; -  
 DR SMART: SM00382; AAA; 1.  
 DR Hypothetical protein; ATP-binding.  
 FT NP\_BIND 72  
 SQ SEQUENCE 465 AA; 51561 MW; 150EB10AD13F847 CRC64;

Query Match 43.7%; Score 989.5; DB 3; Length 465;  
 Best Local Similarity 46.9%; Pred. No. 2.6e-49;  
 Matches 210; Conservative 87; Mismatches 136; Indels 15; Gaps 5;

OY 10 SKRORATHTHRIKGLIDANGMALALAGFVGQAAREAGLAVDMIRQKMAGRAVLLA 69  
 DB 12 SKLERIGASHIRIGIGINDMIEPKESSQGMVGQKARRAGVILKMIQEGRIAGRAIIMA 71  
 OY 70 GPATGKTALALGIAOELGSKVPCPMGSEVYSEVKKTEVLMENFRRAIGRIKENKE 129  
 DB 72 GPSTGKTALAMGASLSDTPVTLASVYSLMSKTEALLQALRSIGRIKETE 131  
 OY 130 YRGEVTELSPEAESTTGGYAKSISHVTSIKTKYKQKOLDSITDALJREKAVGD 189  
 DB 132 IIEGEVVEVQID--RSITGKNKG---KLIRSTDMETVYDGLGFMIDSLRKEVYAGD 185  
 OY 190 VIYEANSAGVAVKRCDFATEYDLEAE--VPIPKGVHKKELVODVTLHDLAAN 247  
 DB 186 VISIDSVGVYVTKLIGSFARARDYDAMGADTRFVOPQOELIKRREVVATSLHDIVN 245  
 OY 248 ADPQGGODILSLMGOMKPRKTEITEKLOEINKVYNNYIDEGIAELVPGVLFIDEVHML 307  
 DB 246 SRQSG--FLAL-----FSDDTGEIKRVEHQITKVSEREGKAEIVPGVLFIDEVHML 298  
 OY 308 DIECSYLRALSPILVILATNGICNVGTDMTSPHGPVLDRLRYIIRFTYGP 367  
 DB 299 DIECSFENRALEDLAPIVAMSNRGTIRGTNYRSPHGPVLDRLRYIIRFTYGP 358  
 OY 368 TMIQILAIRAOVEELIDMEESLAYLGIGQOTSLRHAIIOLSPASVSKTNREKICKA 427  
 DB 359 EVKELIKRCEDEVDMEPSALDYLSTIGQETSLRVALIISNOVALRKSATTEES 418  
 OY 428 DLEEVSGYLDKSSARLLQEOERYIT 455  
 DB 419 DIRRYELFLDOKRSVEVLEEGKNYIT 446

RESULT 11  
 OY 09DE27 PRELIMINARY: PRT; 462 AA.  
 AC 09DE27;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 GN REPTIN.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 OX NCBI\_TaxID=8555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Etard C., Wedlich D., Kuehl M.;  
 RT "Xenopus laevis reptin, a Xenopus homolog of the prokaryotic RuvB  
 RT helicase.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF218071; AAC44126.1; -





RESULT 13			
ID	09V1R4	PRELIMINARY;	PRF; 441 AA.
AC	09V1R4;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2001 (TREMBLrel. 16, Last annotation update)		
DE	TBP-INTERACTING PROTEIN TIP49.		
GN	PAB2107..		
OS	Pyrococcus abyssi.		
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.		
OX	NCBI_TaxID=29292;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ORSAY;		
RA	Hellig R.;		
RT	"Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."		
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ248284; CAB9285.1; -.		
DR	InterPro: IPR000862; -.		
DR	InterPro: IPR001939; -.		
DR	InterPro: IPR003593; -.		
DR	Pfam; PF00004; AAA; 1.		
DR	SMART; SM00362; AAA; 1.		
QO	SEQUENCE 441 AA; 49132 MW; AD733EE96209DDC6 CRC64;		

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Goshl T., Simada M., Kawahire S., Ichimura T., Omata S., Horigane T.;  
 RT "Molecular cloning of mouse p47, a second group mammalian RuvB DNA  
 RT helicase-like protein: homology with those from human and  
 RT Saccharomyces cerevisiae.";  
 RL J. Biochem. 0:0-0(1999).  
 DR EMBL; AB013912; BAA/6297.1;  
 DR MGD; MGI:1342299; Ruvb12.  
 DR InterPro; IPR003593;  
 DR SMART; SM00382; AAA; 1.  
 KM Helicase.  
 SQ SEQUENCE 463 AA; 5112 MM; 419D779972933308 CRC64;

Query Match 43.0%; Score 974; DB 11; Length 463;  
 Best Local Similarity 44.0%; Pred. No. 2e-48;

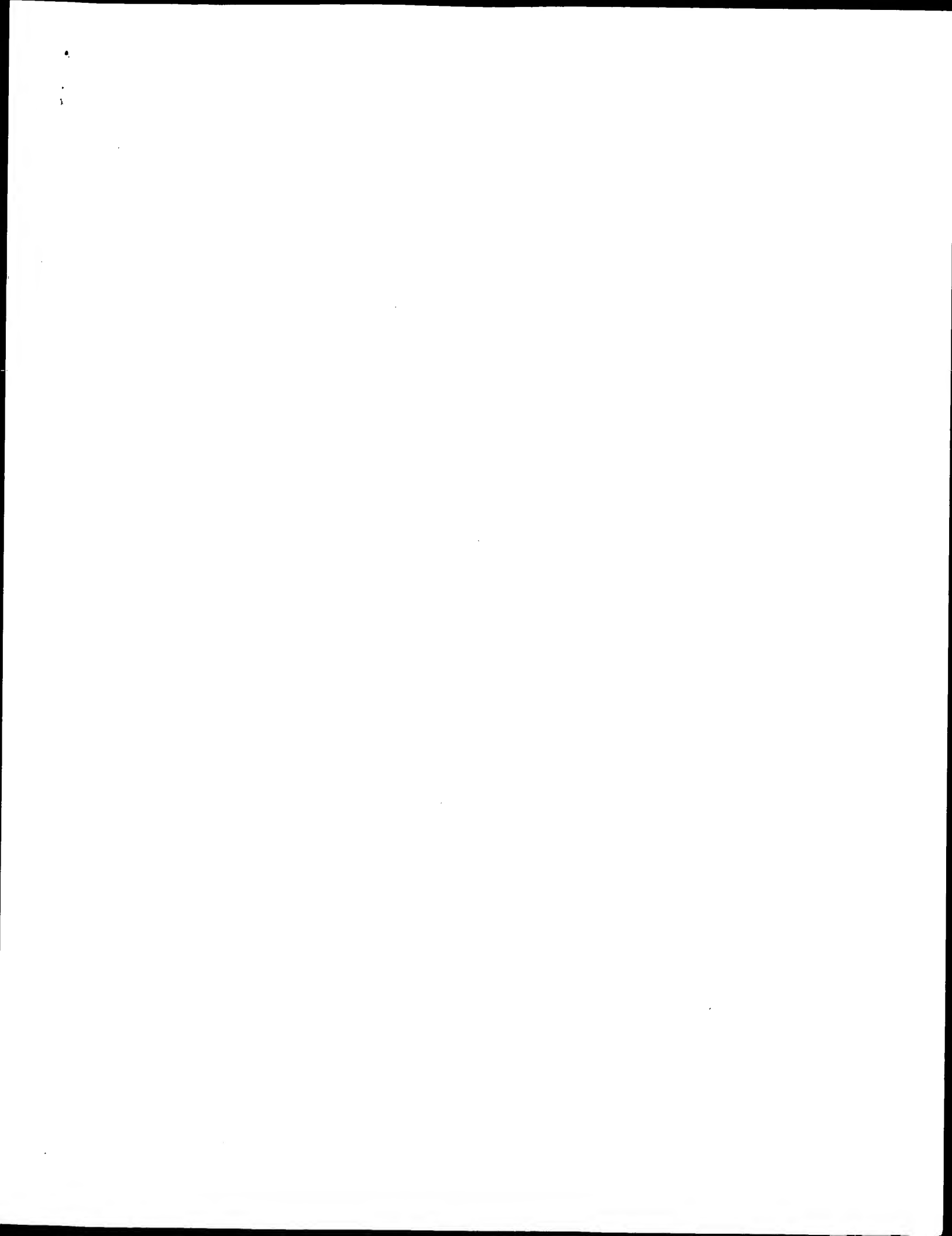
Matches 200; Conservative 97; Mismatches 144; Indels 14; Gaps 4;

```

OY 2 RIEVQSTSKQRIATHTTHIKGLIDANGMAIALAGFVGQAAREAGLAVDMIRQKM 61
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 9 KPEIRIDVTIERIGASHIRGLDLEPRQASQGWQLAARRAGVLEMIREGKI 68
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 62 AGRVLLAGPATKKTALALGIAELGSKVPFCPMVGSSEVYSSEVKTLEMLNFRRAIG 121
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 69 AGRVLLAGPATKKTALALGIAELGSKVPFCPMVGSSEVYSSEVKTLEMLNFRRAIG 128
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 122 LRIKENKEVEEGETELSPFEAEESTTGGYAKSISHVITSLKTVKTKQLKLDSSITDALI 181
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 129 VRIKETETIEEGEVEIQLDRPATGTGSKVKG---LTKTTEMETIYDLGTMIESLT 183
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 182 KERKAVGVDTIYIANSQAVKRYGRCDSPATEYDLAE--YVPIPKGEVHKKEIVODVT 239
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 184 KQVVOAGDVITIDKAKIGISKLGSRFTARADYDAMGSOQKFPQCPDGELOKKEVYHTVS 243
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 240 LHDLDANAQPOGQDILSLMGQMMKPKRTEITEKIROEINKVNVRYIDEGIAELVPGVL 299
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 244 LHEIDVINSRTQG--FLAL---FSGDTGEIKSEVREQINAKVAEWREGEKAEITPGVL 296
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 300 FIDEVHMLDIECESYINRALESPLSIYIATLNRGICNVRGTDMTSPHGIPVDLDRLLVI 359
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 297 FIDEVHMLDIESFSLNRALESMDAPVLLIMATNRGITRIRGTSYOSPHGIPIDLDRLLI 356
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 360 IRTETYGPTMIOQLAIRAOVEIDMDEESLAYIGEIGQOISLRHAQLDISPASVYSKTN 419
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 357 VSTSPYSEKDTKOILIRCEEEVEVMSDEDAYIVLRIGLETSLRAYAQLITAAASLYCKRR 416
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
OY 420 GREKICKADLEEVYSGLYLDKASSARLLQEQERYI 454
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
DB 417 KGEVOVDIKRIVYSLFLDESRSSTQYMKERYODAFI 451
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

Search completed: November 13, 2001, 07:02:32  
 Job time: 353 sec



OM of: US-09-589-510-4 to: Genembi: \* out\_format: pfs

Date: Nov 13, 2001 7:53 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

# Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp  
-O/cgn2.1/USPTO\_pool/US09589510/runat\_13112001\_065634\_13448/app\_query.fasta.1.518  
-DB=Genembi -OFMT=fastap -SUFFIX=std.rge -GAPOP=12.000  
-GAPOP=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -CGAPOP=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPOP=6.000 -FGAPOP=7.000 -START=1 -MATH=10.000 -MATH2  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATH=10.000 -MATH2  
-TRN=human40.cdi -LIST=45 -DOCLIGN=200 -TRN\_SCORE=pt  
-TRN\_MAX=100 -TRN\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09589510\_@cgn1.1.5631 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-589-510-4  
Query length: 455  
Database: Genembi: \*  
Database sequences: 1344157  
Database length: -85606004  
Search time (sec): 1585.400000

## Sequence list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat2:AX077226	+ 2263.00	2375.14	4.6e-124	1912	AX077226 Sequence 3 from Patent
gb_pat2:AX077224	+ 2253.00	2364.97	1.7e-123	1845	AX077224 Sequence 1 from Patent
gb_pat2:AX077228	+ 2222.00	2332.15	1.1e-121	1886	AX077228 Sequence 5 from Patent
gb_pat2:AX077230	+ 2211.50	2321.18	4.7e-121	1898	AX077230 Sequence 7 from Patent
gb_pat2:AX077232	+ 2157.50	2264.66	6.5e-118	1869	AX077232 Sequence 9 from Patent
gb_pat2:AX077234	+ 2039.00	2140.90	1.4e-111	1755	AJ276264 Citer aietinum mRNA
gb_pat2:AX077236	+ 1723.50	1810.01	1.4e-92	1730	E21816 Protein forming a comp
gb_pat2:AX077238	+ 1723.50	1809.98	1.4e-92	1733	AJ010058 Homo sapiens mRNA for
gb_pat2:AX077240	+ 1723.50	1809.93	1.4e-92	1736	AF090984 Homo sapiens Pontin
gb_pat2:AX077242	+ 1723.50	1809.91	1.4e-92	1746	BC002993 Homo sapiens RuvB-1
gb_pat2:AX077244	+ 1723.50	1809.86	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077246	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077248	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077250	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077252	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077254	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077256	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077258	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077260	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077262	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077264	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077266	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077268	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077270	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077272	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077274	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077276	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077278	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077280	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077282	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077284	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077286	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077288	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077290	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077292	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077294	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077296	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077298	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077300	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077302	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077304	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077306	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077308	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077310	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077312	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077314	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077316	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077318	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077320	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077322	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077324	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077326	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077328	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077330	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077332	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077334	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077336	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077338	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077340	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077342	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077344	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077346	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077348	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077350	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077352	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077354	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077356	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077358	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077360	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077362	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077364	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077366	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077368	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077370	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077372	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077374	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077376	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077378	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077380	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077382	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077384	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077386	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077388	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077390	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077392	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077394	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077396	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077398	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077400	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077402	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077404	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077406	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077408	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077410	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077412	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077414	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077416	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077418	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077420	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077422	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077424	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077426	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077428	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077430	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077432	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077434	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077436	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077438	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077440	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077442	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077444	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077446	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077448	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077450	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077452	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077454	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077456	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077458	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077460	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077462	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077464	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077466	+ 1723.50	1809.75	1.4e-92	1750	AF070735 Homo sapiens mRNA for
gb_pat2:AX077468	+ 1723.50	1809.75	1.4e-		

```

|||||
294 CCTGGCGGTCGCCGCCACGGCAAGCGCGCTAGCGCTGGCATG 343
84 laaglulserlyserlyvalprophocysermetvalglyserglu 100
344 CCCAGGAGCTCGCGAGAGGTCCTTTCTGCTATGTTAGATCAGAA 393
101 valtyrsersergluvalylslythrghluvalleumecgluasphar 117
394 GTGACTCTCTCGAGAGTCAGAAAGTGAAGTGTGATGAAAATTTCCG 443
117 gtagalaleglyleuargllelyscluasnlysgluvaltyrghlyg 134
444 TAGAGCTATAGGTTTGGCTATAGAGAAACAAAGAGTTTATGAAGAG 493
134 luvalthrghlyserproglugluvalglulserthrghlyglytyr 150
494 AGGTACTGACATTTCCCGAAGAGGCTGAGAGTACAACTGGTGATAT 543
151 Alalyserlleserhisvallelleserleutystrhrvallysglyth 167
544 GCAAAAGCATTAAGCATTAATCATCATTAAGACTGTAAAGGAGC 593
167 tysglnleuylsleuaspserserlletyraspalaleuileysglu 184
594 TAAGCACTGAGAGTTAGATTTCTCAATTATGATGCTGTATCAAGAAA 643
184 ysvalalavalglaspvalyletyrilleglualaasserglvalval 200
644 AGGTGGCAGTGGGTGATGTTATATCATCATGAGCAAAATAGTGGAGCTG 693
201 lysarvalgltyrghlyserphalathrghlytyraspleugluval 217
694 AAAAGAGTTGGTATGATGATTTCTTTCACAAATACGATCTTGAAGC 743
217 aglulutyrvalproleprolysglygluvalhislyslysglytyr 234
744 TGAAGAGTATGTTCTTATCCCAAGGTGAAGTCCATTAAGAAAGAAA 793
234 levalglaspvalthrleuhalaspaleuasalaasnaaglupro 250
794 TTGTCAGAGATGCACATTCATACCTGATGAGCAAAATGCTCAGCCA 843
251 glnglygluaspilleuuserleumecglylmetketyrproar 267
844 CAAGTGGCCAGATATTTTCTTATGCGCAGAGTGTGTAACCCAG 893
267 glythrghlylethrghlylulysleuarglngluileasnlyvalval 284
894 AAAAGCTGAATCACCGAAAACCTACGCCAAGAAATTAAGTGTGTA 943
284 snatgtyrileaspgluglyllealgluleuvalproglyvalleuphe 300
944 ATGATATATCATGAGAGGATGAGAGCTGTGTACTGCTGTGTTTGTTC 993
301 llaaspgluvalhismetleuaspileglucyspheserlytleuasnar 317
994 ATTGATGAGTCCACATGTTGATATGATGATTTTCTTATCTTAAACCG 1043
317 galaleuylserproleuserproilleuvalleuvalathrasnar 334
1044 TGCATTGGAGAGCCCATTAATCAACATCGTATCTGTATCAAAATAGG 1093
334 lyllecysasnvalargglythrasmethrserprohisglylepro 350
1094 GAATATGATATGTAAGAGGAGTATGACAAAGTCCACATGCTATACCG 1143
351 valaspaleuaspargluvalilleleargthrghlytyrghlytyr 367
1144 GTGGATCTTCTAGATAGGCTGTGATATGTCGAGAGAGACATATGGCC 1193
367 othrghlymetlleglulleuvalleargalaglvalglululea 384
|||||

```

```

1194 TACTGACATGATACAGATATTGGCTATCCGACACAAAGTGAGAGATTG 1243
384 spmetaspglugluserleuvalatyrleuglylulileglylglulthr 400
1244 ATATGATGTAGAGAAAGTCTTTCCTTATTTAGCGAATGCGAGACAGACA 1293
401 Serleuarghisalalegluleuileserproalaservalaserly 417
1294 TCTTTAGACATGCTATTCATTTGATATCACCCTGCCAGCTGTCTCAAA 1343
417 strasnslargglulysllecylysalaaspleugluvalserg 434
1344 GACTTAATGAGAGAGAAATTCGACAGCTGATCTCGAGAGATCAGTG 1393
434 lyleuylrleuaspalalysserseralargleuileuglngln 450
1394 GGCTATATTGGATGCCAAATCTCGCTGGCTGCTCAGAGACACAA 1443
451 GluArgTyrIleThr 455
1444 GAAAGATACATCAC 1458

```

seq\_name: gb\_pat2:AX077224

seq\_documentation\_block:

LOCUS AX077224 1845 bp DNA PAT 22-FEB-2001

DEFINITION Sequence 1 from Patent WO0105975.

ACCESSION AX077224

VERSION AX077224.1 GI:13121818

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

FEATURES

source

CDS

ORIGIN

BASE COUNT

530 a 378 c 490 g 445 t 2 others

alignment\_scores:

Quality: 2253.00

Ratio: 4.952

Percent Similarity: 100.000

Percent Identity: 99.341

alignment\_block:

US-09-589-510-4 x AX077224

Align seg 1/1 to: AX077224 from: 1 to: 1845

1 Metargilleglulvalglulserthrserlysglnargllealath 17

```

|||||
85 ATGAGGATCGAGAGGTGCTCACTCGAATCGAAGAACGACGCGCATCGGCAC 134
|||||
17 THISTHRTISILEUYSGLILEUGLYLEUASPALIAASNGLYMETALILEA 34
|||||
135 CCACACCCACATCAAGGAGCTCGGCTCGACGCCAATGGCGATGGCGATTG 184
|||||
34 IALEUALAIAAGLYPHEVALIGLYGNAIAIAIAIARGGLUAIALAGLY 50
185 CCGTGGCGGGCGGCTCGTGGCGCCAGTGGCGCGCGGAGCGCGCGCGG 234
|||||
51 LEUALAIAASPMEILIEARGINLYSLYMETALAGLYARGALVALLE 67
|||||
235 CTGGCGGCTGCACATGATTGCGCAGAAAAAATGGCGCGCGCGGCTGCT 284
|||||
67 ULLEUALAGLYPROPOIALTHRTGLYLYSTHRLALEUALALEUGLYLEA 84
|||||
285 CCTTGTGGGTCCGCCGCCACGCGCAAGACGCGCTAGCGCTCGGCATAG 334
|||||
84 LAGINLULEUGLYSERLYSVALIPROPHECYSPPROMETVALIGLYSERGLU 100
|||||
335 CCCAGAGAGCTCGGACAGCAAGGTCCCTTCTGCGCTATGTGTAGATCAGAA 384
|||||
101 VALTYRSESERGLIYALLYSLYSTHRTGLUVALLEUMETGLUASNPHEAR 117
|||||
385 GTGTACTCCTCGGAGGTCAAGAAAACTGAGTGTGATGGAATAATTTCCG 434
|||||
117 GATGALIAILEGLYLEUARGLILEUGSLUASNLYSGLIYVALITYRGLUGLYG 134
|||||
435 TAGAGGCTATAGGTTTCCGTATTAAGAAACAAAGAGGTTTATGAAAGGAG 484
|||||
134 IYUVALTHRTGLUSERPROGLUGLUAGLUAGLUSERTHRTHRTGLYLYTYR 150
|||||
485 AGGTTACTGAACTTCCCAAGAGAGGCTGAGTACAACTGCTGGATAT 534
|||||
151 ALALYSSERTLESERHISVALILEISERLEULYSTHRTVALIGLYTH 167
|||||
535 GCAAAAAGCATTTAGCCATGTATATACATTAAACCAATATGAGACAGTG 584
|||||
167 FLYSGINLEULYSLEUASPSERISERTILEYRASPALALEULILEYSGLU 184
|||||
585 TAAACCACTGAAGTTAGATTCTTCAATTTATGATGCTCGATCAAGGAA 634
|||||
184 YSVALIAIAIAGLYASPVAILLETYRILEGLUALAASNSERGLYALVAL 200
635 AGGTGGCAGTGGGTGATTTATATACATTAAACCAATATGAGACAGTG 684
|||||
201 LYSARTYVALIGLYARGCYSASPSPERPHALATHRTGLUTYRSPLEUGLU 217
|||||
685 AAAAGAGTTGGTAGATGTGATCTTTTGTCTACAGAAATACGATCTTGACG 734
|||||
217 AGUGLUYTYRVALPROILEPROLYSGLYLUVALHLSLYSLYSGLU 234
735 TGAAGAAATATGTTCTTATCCCAAGGTGAAGTCCATTAACAAAAAAGAA 784
|||||
234 ILEVALINASPVAILTHRTLEUHSASPLEUASPVAILAIAANALAGLPRO 250
785 TAGTCAGAGATGTCACACTCATGACCTGTGATGAGCAAAATGCTCAGCCA 834
|||||
251 GINGLYGLYGLINASPILLEUSERLEUMETGLYGLIMETMETLYSPROAR 267
835 CAAGGTGGCCAAAGATTTTGTCCCTTATGGCCAGATGATGAACACAGC 884
|||||
267 GLYSTHRTGLUILETHRTGLYLYSLEUARGINLUILIASNLYSVALYALA 284
885 AAAGACTGAATACCCGAAAAAATAAGCCAGAAATTAATTAAGGTGTGATA 934
284 SNAATYTRILEASPGIUGLYILEAGLULEUVALIPROGLYVALLEUPHE 300
935 ATAGATATATGATGAAGGAATGCGAGCTTGTACCTGGTGTATTC 984
301 ILEASPGIYUVALHISMETLEUASPILEGLUCYSPHESERTYRLEUASPAR 317
|||||

```

```

985 ATTGACAGGTCCACATGTTGGATATCGAATGTTTCTTATCTTAACCG 1034
|||||
317 GALALEUGLUSERPROLEUSERPROILEVALILELEUALITHRASPARG 334
|||||
1035 TGCAATTGGAGAGCCCATATACCAATTTGATCTTGTACAGAAATAGGG 1084
|||||
334 IYILECYSASNVALARGLYTHRASPMEITHRSPROHISGLYILEPRO 350
1085 GAATATGTAATCTTAGAGCAACTGATATGACAACTCCACATGATATMCA 1134
|||||
351 VALASPLEUASPARLEUVALILEIARGYTHRTGLUTYRGLYPR 367
1135 GTGGATCTTCTGATAGGTGGTGTATTTTCGACAGAGACATATGCGCC 1184
|||||
367 OTHRTGLMETILEGINILELEUALIAIIEARGALAGIYAGLUJULILEA 384
1185 TACTGAGATGATACAGATATGCTATCCGAGCACAGATGAGACATTTG 1234
|||||
384 SPMEIASPGIUGLUSERLEUALIATYRLEUGLYGLIILEGLYGLINTHR 400
1235 ATRTGATACAGAAAGTCTGCTTATTTAAGGAGATCGACAGACAGACA 1284
|||||
401 SERLEUARGHISALALEGINLEULIESERPROIALASERVALIASERTLY 417
1285 TCTTTAAGACATGCTATTCATTTGATATCACCTGCCACGCTGCTCAA 1334
|||||
417 STRHANSGLYARGGLIYSLILECYSLSYSLASPLEUGLUVALISERG 434
1335 GACTAATGGAAGAGAGAAATATGCAAGGCTGATCTCGAGAAAGTCACTG 1384
|||||
434 IYLEUTYRLEUASVALIYSSERIALIARGLEULEUGLUGLUGLIN 450
1385 GGCTCATTTTGGATGGCAATCTCGCTGCGCTGCTCGACAGACAGACA 1434
|||||
451 GLUARGTYRILETHR 455
1435 GAAAGATACATCAC 1449
seq_name: gb_pat2:AX077228
seq_documentation_block: 1886 bp DNA PAT 22-FEB-2001
LOCUS AX077228 5 from Patent WO0105975.
DEFINITION Sequence 5 from Patent WO0105975.
ACCESSION AX077228
VERSION AX077228.1 GI:13121822
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 1886)
Mahajan, P.B.
Maize orthologues of bacterial ruvB:cdnas and uses thereof
Patent: WO 0105975-A 5 25-JAN-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
location/Qualifiers
1..1886
/organism="Zea mays"
/db_xref="taxon:4577"
82..1449
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC32390.1"
/db_xref="GI:13121823"
/translat="MRIEVOSTSKKRIATHHTIKIGLDANCKSNPLAAGFYGOAA
AREAGIADVMIROKMAGRALILAGPATCTTALALIAELSLKVPFCPMSSEY
SSPVRKTEVMEFRRAIGIRIKENKEVEGEVILESPAESETGGYAKAISVITIG
LKTQCTLOKLIDPSITYDALIKENVAAGVITYEANSAGAVRVGCSFAEYDLEA
EYPRKEGVHNKKELIYODVTHLADIAAAGOGDILSLMGQMKRKEITEKLR
OETNKVNNRTIDBGLAELVPGVLEFDENYHMLDEGFSYLNRLASPLSYIATNRG
ICNVRGDMTSPHGIPVDLLDRVILIRETTYPTPTMOTIALRQVEYIDIDESLAY
LGITGOOTSIRHAIQDLSPASVAVAKTNGEKKKADLEVSGLYIDAKSSARLLQEOQ

```

BASE COUNT 524 a 408 c 518 g 436 t  
 ORIGIN  
 Alignment scores:  
 Quality: 2222.00 Length: 455  
 Ratio: 4.916 Caps: 0  
 Percent Similarity: 99.341 Percent Identity: 97.582  
 alignment\_block:  
 US-09-589-510-4 x AX077228 ..  
 Align seg 1/1 to: AX077228 from: 1 to: 1886

```

1 MetArgIleGluGluValGlnSerThrSerLysLysGluArgIleAlaTh 17
  |||||||
82 ATGAGATGAGAGAGGTGCATCGACTCGAAGAAAGCAGCGCATCGCCAC 131
  |||||||
17 rHsThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
  |||||||
132 CCACACTCACATCAAGGCGCTCGCTCGACGCCAATGGAATGTGATCC 181
  |||||||
34 LalaValAlaGlyPheValGlyValAlaAlaArgGluAlaAlaGly 50
  |||||||
182 CGTTGGCGGGGGGCTTCGTGGCCAGGCGGGCGCGGAGCGCGCCGCG 231
  |||||||
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaValle 67
  |||||||
232 CTGGCGCTGCATGATCGCGCAGAAAGATGGCGGCTCGCGCTGCT 281
  |||||||
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyLea 84
  |||||||
282 CTTTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 331
  |||||||
84 laGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100
  |||||||
332 CCGAGAGAGCTCGGCGAGCAAGGCTCCATCTCTCTATGCTAGATCGAA 381
  |||||||
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
  |||||||
382 GGTACTCTCTCAGAGGTCAGAAACAGAGGTGCTGATGGAATATTCGC 431
  |||||||
117 GARAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134
  |||||||
432 TAGAGCTTAGGTTTGGCTATTAAGAAACCAAGAGGTTATGAAGGAG 481
  |||||||
134 luValThrGluLeuSerProGluGluAlaGlnSerThrThrGlyGlyTyr 150
  |||||||
482 AGTTATGAGACTTTCCTCCAGAAAGGCTGAGAGCAACACTGGTGGATAT 531
  |||||||
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167
  |||||||
532 GCGAAAGCATTTAGCCACGTAATCATTTGGCTTAAAGACTGTCAAGGAG 581
  |||||||
167 rLysGluLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
  |||||||
582 TAAGCAATTGAANTAGACCCCTCATTTATGATGCTTGTATCAAGGAA 631
  |||||||
184 yValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
  |||||||
632 AGGGGGAGAGGGGTGATGTTATATACATGAAGCAAAATAGTGAGCAGTG 681
  |||||||
201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217
  |||||||
682 AAAAGAGTTGGTAGATGATCTTTTGGCTACAGATATGATCTTGAAGC 731
  |||||||
217 agLugLysTyrValProIleProLysGlyGluValHisLysLysGlyLut 234
  |||||||
732 TGAAGAGATAGTCTATCCCAAGGGAATCCATTAAGAAAAGGAAA 781
  |||||||
234 lValAlaLysPvalThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
  |||||||

```

```

782 TAGTGCAGAGATGTACACTTCATGACCTTGATGACAGCAATGCCAGCA 831
  |||||||
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267
  |||||||
832 CAAGGTGGCCAGATATTTTGTCTTATGGCCAGATGATGATGACGACG 881
  |||||||
267 GlyThrGluIleThrGluLysLeuArgGlnLutIleAsnLysValAla 284
  |||||||
882 GAAGACTGAATTCACCCGAAAGCTACGCCAAGAAATCATAGTGGTAA 931
  |||||||
284 snArgTyrIleAspGlnGlyIleAlaGluLeuValProGlyValLeuPhe 300
  |||||||
932 ACAGATATATCGACGAGAGATTCGAGAGCTTGATGCTGGTGTGTTTC 981
  |||||||
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
  |||||||
982 ATTGATGAGAGTCCACATGTTGATATTGAAATGCTTTTCTTATTAACG 1031
  |||||||
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
  |||||||
1032 TGCATTGGAGAGCCCATATACCAATTTGTATCTCCTACGAATAGAG 1081
  |||||||
334 lYlLeCysAsnValArgGlyThrAspMetThrSerProHisGlyIlePro 350
  |||||||
1082 GATATATGTAATGTGAGAGAACCGATATGACGAGTCCATGTATACCA 1131
  |||||||
351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPr 367
  |||||||
1132 GTGGACCTTTAGATAGCTTGGTGAATATTCGACAGAAACATATGCCC 1181
  |||||||
367 cThrGluMetIleGlnIleLeuAlaIleArgAlaGluValGluGluIleA 384
  |||||||
1182 TACTGAGATGATACGATACGATGCTGCTATCCGACCAAGTGAAGAGATTG 1231
  |||||||
384 sPMeAspGluGluSerLeuAlaTyrLeuGlyLutIleGlyGlnIleThr 400
  |||||||
1232 ATATCGATGAAGAAAGCTTCTTATTTAGCGAGATCGGACGACGACA 1281
  |||||||
401 SerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValValSerLy 417
  |||||||
1282 TCTTTGACACATGCTATTTCATTCATCCTCCAGCGCTGCTGCCAAA 1331
  |||||||
417 sThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluGluValSerG 434
  |||||||
1332 GACCAAGCGGAGAGAAAGATGTCCAAAGGCTGACCTCGAGGAAAGTCAAG 1381
  |||||||
434 lYLeuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlnGln 450
  |||||||
1382 GGCTCTATTTTGGATGCCAATCTCGGCTGCTGCTCCAGAGGACACNA 1431
  |||||||
451 GluArgTyrIlePhe 455
  |||||||
1432 GAAAGATACATCAC 1446
  |||||||
seq_name: gb_pat2:AX077230
seq_documentation_block:
LOCUS AX077230 1898 bp DNA PAT 22-FEB-2001
DEFINITION Sequence 7 from Patent WO0105975.
ACCESSION AX077230
VERSION AX077230.1 GI:13121824
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 1898)
AUTHORS Mahajan,P.B.
TITLE Maize orthologues of bacterial ruvB:cdnas and uses thereof
JOURNAL Patent: WO 0105975-A 7 25-JAN-2001.
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers

```



```

source
1. 1898
/organism="Zea mays"
/db_xref="taxon:4577"
166. 1536
/note="unnamed protein product"
CDS
/codon_start=1
/protein_id="CAC32391.1"
/db_xref="GI:13121825"
/translation="MREIVQSTSKKRIATHTHTIKGLIDQANGSMPLAFGEVGA
AAREAGLAVDMIRKRMGRALLLAGPRATKALAGIAQELGSKVPCPOMVSEV
YSSEVKTEVLEMFERRAIGLRKENKEVEEVEIEENSAVRRCDSFRTEYDEA
GLTKVGTAKLIDPSIDALIKEKVAADVITYEENSAVRRCDSFRTEYDEA
EEVYPIKPEVHKKEIVODVTLHEDANADPOGGDILSLMGMMKPKTEITKL
ROENKVNREYIDEGIAELVPGVLEFEDVHMDIECFYLNRLSPSPVILATNR
GICNVGTQMTSPHGIIPVDLDRVLIIRTEGPTEMIOLAIRAVEEIDIDESLA
YLGIRGQTSRLRAIQLLSASVAVATNREKMKRADLEVSGLYLDKSSARLIQEQ
QERTY"
BASE COUNT 532 a 413 c 521 g 432 t
ORIGIN

alignment_scores:
quality: 2211.50 Length: 456
Ratio: 4.893 Gaps: 1
Percent Similarity: 99.123 Percent Identity: 97.368

alignment_block:
US-09-589-510-4 x AX077230 ..
Align seq 1/1 to: AX077230 from: 1 to: 1898

1 Metatrgllleugluvalginserthrseryslysglaarglileath 17
166 ATGAGGATAGAGGAGGTGCAATCGACCTCGAAGAACACGCGATCGCCAC 215
17 RHSTHRIHISILYSGLYLEUENLYLEUASP...Alaasnglymetala 33
216 CCACACTCATCATCAAGGCGCTCGCGCTCGACGCGCATTCGATGTGCA 265
33 lealaleualaialaglyphevalglylnalaialaalaargluuala 49
266 TGCCCTTGCGGCGCGGCTGCTGCGCGAGCGCGCGCGCGCGCGCGC 315
50 Glyleualaialaspmetilearlglnlyslslysmetalaargla 66
316 GGGGTGGGCGTGCACATGATCCGCGACAGAACAGATGCGCGTGCCTG 365
66 lleuleualaialaglyprothalaithrglylsthralaleuala 83
366 GCTCCTTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 415
83 lealaglnleuglyserlyvalprophecysprometvalglyser 99
416 TAGCGCAGACGCTCGCGACGCAAGGTCCCATTCGTCTATGTAGATCA 465
100 Gluvaltyrsersegluvallyslsthrgluvalleumetgluansp 116
466 GAAGGTACTCTCTCAGAGTCAGAAACAGAGTGTGATGGAATTTT 515
116 eargtrglaialleuglyleualaarglyleuasnlysgluvaltyr 133
516 CCGTAGACGTATAGGTTGCGTATTAAGAAACAAAGAGTTATCAAG 565
133 lygluvalthrgluuenserprogluualagluuserththrglygly 149
566 GAGAGGTATGTAACCTTCCAGAGAGGCTGAGAGCACAACTGGGTGA 615
150 tyrralalysertleaserhisvalleleaserleuysthrvalysg 166
616 TATGCGAAAGCATTAAGCACCTATTCATTCGTTAAAGACTGTCAAG 665
166 ythrlysglnleuysleuaspserserlethyaspalaileuileysg 183

```

```

666 GACTAAGCAATTGAATTAAGACCTTCATTTATGATGCTTGTATCAAG 715
183 Iulysvalalaialaspvallethyrlleuglialaasnserglyala 199
716 AAAAGTGGCAGTGGTGATGTTATATACATGAAGCAATATAGGAGCA 765
200 vallysarvalglyarqcyaspsersphealathrgluuyspleugl 216
216 valagluuvalthrgluuenserprolysglylnalaialaasnlysg 233
816 ACTCGAAGGTATGTTCTTATCCCAAGGTGAAGTCCATTAAGAAAG 865
233 Iutlevalglnaspvalthrlleuhsaspleuaspaalaalasnalgln 249
866 AATATGTGAGATGTACACTCCATGACCTTGATGCGCAAAAGCCAG 915
250 Proglnglyglylnaspilleuenserleumetglylnmetleuyspr 266
916 CCACAAAGTGGCCAGATATTTTGTCCCTTATGCGCAGATGATGAAGCC 965
266 oarglysthrgluuethrgluuysleuarglnluileasnlysalv 283
966 ACAGAACTGAATCAACCAAGCTACGCGCAAGAAATCAATAGGTGG 1015
283 alasnarglyrileaspuglylilealagluvalproglyvalleu 299
1016 TAAACAGATATATCGACGAAAGATGCGAGACTTGCTACCTGCTTGTG 1065
300 Pheileaspgluvalhismetleuasplleuglycyspheserlyleuas 316
1066 TTCATTGATGAGGTCCACATGTTGATATGTAATGCTTTTCTTATCTTAA 1115
316 nargalaleugluuserproleuserprollevalleleualaithrasna 333
1116 CCGTGCTTGAGAGGCCATTATACCAATGTGATACCTGGTACAGATA 1165
333 rglglylecyasnvalargglythrAspmethrserprohisglyle 349
1166 GAGGAATATATATGTAGAGAAACGATATGACGATCCACATGGTATA 1215
350 Provalaspleuasnaspargleuvalleleargthrgluuethrtyr 366
1216 CCAGTGGACCTTCTAGATAGGTTGGTCAATATTCGAGCAAGAAACATATGG 1265
366 yprothrgluuethrleuglnleualaialeargalaglnvalglu 383
1266 CCCTACTGAGATGATACAGATACGTGCTATCCGAGCACAAGTGAAGAGA 1315
383 leaspmetaspgluuserleuualathrleuglyluileglylngln 399
1316 TTGATATCATGAGAAAGAGTCTGCTTATTTAGCGGAGCGAGACGACG 1365
400 Thrsleuathrglnsalaleuglnleuileserproalaservalys 416
1366 ACATCTTTGAGACATGCTATTCAGTTGCTATACCTCCAGCGGTGGCGC 4415
416 rlysthrasnlglyarqgluylslecylysalaaspleuugluvalys 433
1416 AAAGACCAAGCGGAGAGAAAGATGTGCAAGCGCTGAGAGAACTCA 1465
433 erglyleuuthrleuaspaalyserserAlaargleuueuglnlgln 449
1466 GCGGCTCTATTTGGATGCCAAATCTCGCTGCTGCTCAGAGACCA 1515
450 Glngluargtyrlethr 455
1516 CAAGAAAGATATACATCACC 1533
seq_name: gb_pat2:AX077232
seq_documentation_block:

```

LOCUS AX077232 1869 bp DNA PAT 22-FEB-2001  
 DEFINITION Sequence 9 from Patent WO0105975.  
 ACCESSION AX077232  
 VERSION AX077232.1 GI:13121826  
 KEYWORDS  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 Panicoidae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 1869)  
 AUTHORS Mahajan, P.B.  
 TITLE Maize orthologues of bacterial ruvB:cdnas and uses thereof  
 JOURNAL Patent: WO 0105975-A 9 25-JAN-2001;  
 PIONEER HI-BRED INTERNATIONAL, INC. (US)  
 FEATURES  
 source  
 1. 1869  
 location/Qualifiers  
 /organism="Zea mays"  
 /db\_xref="taxon:4577"  
 64. 11380  
 /note="unnamed protein product"  
 /codon\_start=-1  
 /protein\_id="CAC32392.1"  
 /db\_xref="GI:13121827"  
 /translation="MREEVQSTSKQRIATTHIKGLDANGNAIALAAGFVGOK  
 MAGRAVLNAGPRATGKTALAGLDELGSVPPVSGSEVSSSEVKTEVLMENFRA  
 IGRIRKENREVEYEGEVELSPBEASTETGVAKSIHV1ISIKTKGKOLKDSIY  
 DALIKERVAAGDVITYEANSQAVIKRVGRCDSPATYEDLEAEYVP1PGVEVKKKEIY  
 ODVTLHDVAANOPOGODILSMGOMKPKRTEIKTLOEINIKVNRXRTDGSIAE  
 LVPGVLFIDEVHMLDIECFESYLNRALESPLSIYIATLNRCINCRGTDNTPSGIIV  
 DLDRLVIRTEYVGTETMIOILATRAOYERDMDEESLALGEGDQTSURHA10LI  
 SPASVSKNGNEKICKADLEVSGLYIDAKSSARLDEQDERVIT"

BASE COUNT 539 a 373 c 486 g 471 t  
 ORIGIN

alignment\_scores:  
 Quality: 2157.50 Length: 455  
 Ratio: 4.926 Gaps: 2  
 Percent Similarity: 96.264 Percent Identity: 96.264

alignment block:  
 US-09-589-510-4 x AX077232 ..

Align seg 1/1 to: AX077232 from: 1 to: 1869

1 MetArglleGluGluValaGlnSerThrSerLysLysGlnArgTLeaLath 17  
 64 ATGAGGATCGAGGAGGTGCGACCTCGAAGAACAGCGCATCGCCAC 113  
 17 ThrThrHisIleLysGlyLeuGlyLeuAspLysPheLysGlnMetLysLys 34  
 114 CCACACCCATCATCAGGAGACTCGGCTCCACGCCAATGGATGGGATGG 163  
 34 LeuLeuAlaIleGlyPheValaGlyGlnAlaAlaAlaArgGlnAlaIleGly 50  
 164 CGTTGGCGGGGGGTTCTGGGGC..... 186  
 51 LeuAlaValAspMetIleArgGlnLysLysMetLysArgAlaValLea 67  
 187 .....CAGAGAGAGATGGCGCGCGCGCTGCT 215  
 67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyLea 84  
 216 CCTTGGGGGTCCGCCCGCCAGCGGCAAGCGCGCTAGC...GGCATAG 262  
 84 IeGluGluLeuGlySerLysValaProPheCysProMetValaGlySerGlu 100  
 263 CCAGAGAGCTCGGAGAGGCTCTTCTCTCTATGTGTAGATCAGAA 312  
 101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheArg 117  
 313 GTTACTCTCTCGAGAGTCAAGAAACTGAGGTCTGATGGAATAATTTCCG 362

117 gATGAlaIleGlyLeuArgIleLysGluAsnLysGluValTyrGluGly 134  
 363 TAGAGCTATAGGTTTCCGTATTAAGAAACAAAGAGGTTTATGAGAGAG 412  
 134 JuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyGlyTyr 150  
 413 AGGTACTAGAACTTTCCCGAAGAGGCTGAGAGTACAACTGGTGATAT 462  
 151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167  
 463 CCAAAAGATTAAGCATTAATCATCACTTAAGAACTGTTAAAGGAC 512  
 167 rLysGluLeuLysLeuAspSerSerIleTyrAspAlaLeuLysGluL 184  
 513 TAAGCACTGAAGTTAGATTCTTCAATTATGATGCTGTGATCAAGAAA 562  
 184 yValAlaValaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200  
 563 AGGTGGCACTGGGTGATGTTATATACATCGAAGCAAAATAGTGGACAGT 612  
 201 LysArgValaGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
 613 AAAAGATTGGTAGATGATGATTTTTCCTACAGAAATACATCTTGAGAC 662  
 217 gLuuGluTyrValaProIleProLysGlyGluValHisLysLysGluL 234  
 663 TGAAGAGATGTCTCTATCCCAAGGTAAGTCCATTAAGAAAGAAA 712  
 234 LeValGluAspValThrLeuHisAspLeuAspLysAlaAlaAsnArgInPro 250  
 713 TTGTGCAGAGATGTCACACTTCATGACCTTATGACGACCAATGCTCAGCA 762  
 251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProArg 267  
 763 CAAGTGCCCAAGATATTTTGTCCCTATATGCGCCAGATGTAAGAACAG 812  
 267 gLysThrGluIleThrGluLysLeuArgGlnGluLysLysValaVala 284  
 813 AAAGACTGAATATCCGAAAGAACTACGCCCAAGAAATTAATTAAGGTGCT 862  
 284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValaLeuPhe 300  
 863 ATGATATATCATCATGAAGAAATGACAGACTGTGACTGTGTTTGTTC 912  
 301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnArg 317  
 913 ATGATGAGGTCCACATGTTGATATGCAATGTTTCTTATCTTAACCG 962  
 317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334  
 963 TGCATTGAGAGGCCATTTATCAACAAATGCTGATCTGTCACAAATTAAG 1012  
 334 LysIleCysAsnValaArgGlyThrAspMetThrSerProHisGlyLeuPro 350  
 1013 GAATATGATATGTAAGAGGAACTGATATGACAAATGTCACATGATATGAC 1062  
 351 ValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrGlyPro 367  
 1063 GTGATCTTCTAGATAGGCTGTGATTTATGTCGACAGAGACATATGCGCC 1112  
 367 oThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluLysLea 384  
 1113 TACTGATGATACAGATATATGCTATCCGACCAAACTGAGAGAGATTG 1162  
 384 spMetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGlnThr 400  
 1163 ATATGATGAAGAAAGCTCTTCTATTTAGCGCAGATCGGACAGACAGACA 1212  
 401 SerLeuAlaGlnHisAlaIleGlnLeuIleSerProLysSerValaSerLys 417  
 1213 TCTTGAACATGCTATTCATTTGATATCACTGACAGCGTGTCTCAAA 1262

417 sThAsnGlyArgGluIuysTleCysLysAlaAspLeuGluValSerG 434  
|||||  
1263 GACTAAAGGAGAGAAAATCTGCAAGCTGATCTCGAGAGAGTTAGTG 1312  
434 LLeuTYrLeuAspAlaLysSerSerAlaArgLeuEngIngluInGln 450  
|||||  
1313 GGCTCTATTGGATGCAATCCCTGCGTCTGCTCCAGAGCAACAA 1362  
451 GLuArgTYrIleThr 455  
|||||  
1363 GAAAGATACATCACC 1377

seq\_name: gb\_D13:CAR276264  
seq\_documentation\_block:  
LOCUS CAR276264 1755 bp mRNA PLN 06-MAR-2000  
DEFINITION Cicer arletinum mRNA for putative Ruv DNA-helicase.  
ACCESSION AJ276264  
VERSION AJ276264.1 GI:7208770  
KEYWORDS putative Ruv DNA-helicase.  
SOURCE chickpea  
ORGANISM Cicer arletinum  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
Fabales; Fabaceae; Papilionoideae; Cicer.  
1 (bases 1 to 1755)  
Dopico, B., Esteban, R. and Labrador, E.  
A RuvB DNA-helicase like protein is expressed in chickpea epicotyls  
Unpublished  
2 (bases 1 to 1755)  
Labrador, E.  
Direct Submission  
Submitted (29-FEB-2000) Labrador E., Dpto. Fisiologia Vegetal,  
Univ. Salamanca, Campus Miguel de Unamuno. Pza. Doctores de la  
Reina s/n, E-37007, SPAIN  
Location/Organism  
1. 1755  
/organism="Cicer arletinum"  
/cultivar="Castellana"  
/db\_xref="taxon:3827"  
/country="Spain"  
/dev\_stage="seeding"  
/tissue\_type="etiolated epicotyls"  
/clone\_lib="CAN-5"  
/clone="CanRuv"  
/note="age 5 days"  
207. 1583  
/codon\_start=1  
/product="putative Ruv DNA-helicase"  
/protein\_id="CAB76908.1"  
/db\_xref="gi:7208771"

FEATURES  
source  
1. 1755  
/organism="Cicer arletinum"  
/cultivar="Castellana"  
/db\_xref="taxon:3827"  
/country="Spain"  
/dev\_stage="seeding"  
/tissue\_type="etiolated epicotyls"  
/clone\_lib="CAN-5"  
/clone="CanRuv"  
/note="age 5 days"  
207. 1583  
/codon\_start=1  
/product="putative Ruv DNA-helicase"  
/protein\_id="CAB76908.1"  
/db\_xref="gi:7208771"

BASE COUNT 525 a 332 c 410 g 488 t  
ORIGIN  
alignment\_scores:  
Quality: 2039.00 Length: 455  
Ratio: 4.582 Gaps: 0  
Percent Similarity: 97.802 Percent Identity: 87.473  
alignment\_block:  
US-09-589-510-4 x CAR276264 ..  
Align seg 1/1 to: CAR276264 from: 1 to: 1755

1 MeTArgIleGluValAlaGlnSerThrSerLysLysGluArgIleAlaTh 17  
|||||  
216 ATGAAATAGAGAGAGTTCAATCCACGACGAGAAACACAGGACTTCTAC 265  
17 rHisThrHisIleLysGlyLeuEngLysLeuAspAlaAsnGlyMetAlaIleA 34  
|||||  
266 TCACACTACACTTAAAGGCTAGGTCTCGAGGTGAGTGAAGAGATTGCG 315  
34 LLeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaAlaGly 50  
|||||  
316 CTTTTCATCTGATTTGTTGGGCGCAAGCTGAGGCAAGAGAGCATGTGGG 365  
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaValle 67  
|||||  
366 CTTGTGTGATATGATATGCTGACAGAGAGATGGCTGCAAGCAGCTCC 415  
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84  
|||||  
416 TCTCGCGGGCCACCTGGGCGCAAGACTGCTTGTAGCACTTGGGATAT 465  
84 LLeuGluLeuEngLysSerLysValProPheCysPrometValGlySerGlu 100  
|||||  
466 GTCAGAGCTTGGGACCAAGGTTCCATTTCGCCAATGTTAGTTTCAGAA 515  
101 ValTYrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117  
|||||  
516 GTATATTTCACAGAAAGTAAAGAGACGAGGTCTTATGAAATTTTCG 565  
117 gArgAlaIleGlyLeuArgIleLysGluAsnLysGluValTYrGluGly 134  
|||||  
566 ACGGGCTATTGGTCTACGATTAAGGAAATTAAGGAAATATATAGAGGCG 615  
134 LuValThrLysLeuSerProGluGluAlaGluSerThrThrGlyGlyTYr 150  
|||||  
616 AGGTGACTGAGCTCTCCCGAAGAACTGAGAGCTTAAACAGAGGTAT 665  
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTh 167  
|||||  
666 GGTAAAGATTATAGCCAGCTGATTAATGGTGAAGAGAGTTAAAGAGAC 715  
167 rLysGlnLeuLysLysLeuAspSerSerIleTYrAspAlaLeuIleLysGlu 184  
|||||  
716 CAAGCAACTAAAGTTGGACCCCAAGATATATATGCTCTTATTAAGGAAA 765  
184 yValAlaValAlaGlyAspValIleTYrIleGluAlaAsnSerGlyAlaVal 200  
|||||  
766 AGGTAGCTGTGGGAGATGTTATATATATAGGCAAAATAGCGGCGTGTG 815  
201 LysArgValAlaArgCysAspSerPheAlaThrGluTYrAspLeuGluAl 217  
|||||  
816 AAAAGGCTTGGCCGAAAGTATGCTTTTGTACTGAGATTTGACCTTGAAAC 865  
217 aGluGluTYrValProIleProLysGlyGluValHisLysLysGlyIle 234  
|||||  
866 AGAGAGATGATGTTCCACTTCTTAAGGAGAGGCTTCACAAAAAAGAGA 915  
234 LeValGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250  
|||||  
916 TTGTTCAAGATGTAAACCTACATCATCTGATCTGCAAGCAGACACT 965  
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267  
|||||  
966 CAAGGGGGCAAGATATTTTGTCTTATGAGGCCAGATGATGAAGCTAG 1015  
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284  
|||||  
1016 GAAAAACAATACACTGACAACTGAGACAAAGAAATTAATTAAGGTTGCTCA 1065  
284 snArgTYrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300  
|||||  
1066 ACCGATATATGATGAAGTGTGACAGACTTGTCTCGGGGTTCTATTT 1115

[illegible]

```

167 rlyscgluLeuLysLeuAspSerSerIleTyrAlaLeuIleLysGluL 184
|||||
554 CAACAGCTTGAACACTGACCCAGCATTTTGAAGTTTGCAAGAAAGAC 603
|||||
184 yValAlaValAlaLysPValIleTyrIleGluAlaAsnSerGlyAlaVal 200
:::|
604 GAGTAGAAGCTGAGATGTGATTTACATTGAAGCCAAACAGTGGGCGCTG 653
|||||
201 LysArgValAlaArgCysAspSerPheAlaTrnGluTyrAspLeuGluL 217
|||||
654 AAGAGGACGAGGCGAGTGTATACCTTATGCCACAGAAATTCGACCTTGAA 703
|||||
217 agluGluTyrValProIleProLysGlyLysValAlaHisLysLysGluI 234
|||||
704 TGAAGAGTATGTCCCTTGCCAAAGCGGATGTCACAAAAAGAAAGAA 753
|||||
234 lValAlaGlnAspValThrLeuHisAspLeuAspAlaAlaAsnIleGlnPro 250
||:|
754 TCATCCAGAGTGTGACCTTGACATGACTTGATGCTGAATGCGCGGCC 803
|||||
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetLysProAr 267
|||||
804 CAGGGGGAACAAAGATATCCTGTCATGATGAGCGCCACCTATGAAGCCAA 853
|||||
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
:::|
854 GAAGACAGAAATCAGACAACTGAGGAGGATTAATGAAGTGGTGA 903
|||||
284 snArgTyrIleAspGlnGlyIleAlaGluLeuValProGlyValLeuPhe 300
||:|
904 ACAAGTACATGACACAGGCAATGCTGAGCTGCTCCGGGTGCTGTT 953
|||||
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
:::|
954 GTTGATGAGGTCACATGCTGACATGATGATGCTTCACTTCACTGACCG 1003
|||||
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
|||||
1004 CGCCCTGAGCTTCTATGCTGCCATGCTATCTTTCATCCAAACCGAG 1053
|||||
334 lYleCysAsnValArgIYThr...AspMetThrSerProHisGlyIle 349
||:|
1054 GCAACTGTGTATCAGAGGACACTGAGACATCACAATCCCTCAAGGATC 1103
|||||
350 ProValAspLeuLeuAspArgLeuValIleIleArgThrGluThrGly 366
||:|
1104 CCTCTGACCTTGTGACCGAGTGTATATCCGAGCATGCTGTATAC 1153
|||||
366 yProThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGlu 383
||:|
1154 TCCACAGGAATGAACAGATCATTAATAATCCGTCAGAGGAGAA 1203
|||||
383 leAspMetAspGluGluSerLeuAlaTyrLeuGlyGluIleGlyGln 399
||:|
1204 TCACATCAGAGGAGGAGGACACTGAACACCTGGGAGATTCGACCAAG 1253
|||||
400 ThrSerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValIle 416
||:|
1254 ACCACACTGAGTACTCAGTGCACCTGCTGACCCCGGCAACTGTGCTGC 1303
|||||
416 fLysThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluValS 433
||:|
1304 TAAATTCACAGGGAAGACAGCATTTGAGAAAGACATCTGCACAGATCA 1353
|||||
433 eTgLYleuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlu 449
||:|
1354 GTGACTTTTCTATGATGCCAAGTCTCCGCCAAATATCTGCTGACAG 1403
|||||
450 GlnGluArgTyrIle 454
|||||
1404 CAGGATTAAGTACATG 1418
|||||
seq_name: gb_pat2:E21816

```

---

```

seq_documentation_block: 1730 bp DNA PAT 07-FEB-2001
LOCUS E21816
DEFINITION Protein forming a complex with TBP, polynucleotide encoding said
protein, antisense polynucleotide against said polynucleotide, and
antibody recognizing said protein.
ACCESSION E21816
VERSION JP 1999009285-A/2.
KEYWORDS JP 1999009285-A/2.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1730)
AUTHORS Toshihiro, K.T.T. and Makino.
TITLE Protein forming a complex with TBP, polynucleotide encoding said
protein, antisense polynucleotide against said polynucleotide, and
antibody recognizing said protein.
JOURNAL Patent: JP 1999009285-A 2 19-JAN-1999;
SUMITOMO ELECTRIC IND LTD
COMMENT OS Unidentified
PN JP 1999009285-A/2
PD 19-JAN-1999
PF 27-JUN-1997 JP 1997187398
PR
PI TOSHIHIRO, KISHIMOTO, TANAOKI TAMURA, YASUTAKA MAKINO PC
C12N15/09,A61K48/00,C07K14/47,C07K16/18,C12P21/02,C12P21/08, PC
C12Q1/68
PC C12N15/00
CC Topology: Linear; Location/Qualifiers
FH Key 1..1730
FT source /organism='unidentified'.
FEATURES
source 1..1730
location/Qualifiers
BASE COUNT 469 a 401 c 476 g 384 t
ORIGIN
alignment_scores:
Quality: 1723.50 Length: 455
Ratio: 4.113 Gaps: 1
Percent Similarity: 92.088 Percent Identity: 72.527
alignment_block:
US-09-589-510-4 x E21816 ..
Align seg 1/1 to: E21816 from: 1 to: 1730
1 MetArgIleGluGluValGlnSerThrSerLysGlnArgIleAlaTrh 17
|||||
60 ATTAAGATTGAGGAGGTGAAGACACTAGCAAGACGACGATGCGCTC 109
|||||
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
:::|
110 CCACAGCCACAGTGAAGAGGCTGGGCTGAGACGACGCGCTGGCCAGC 159
|||||
34 lAlaLeuAlaIleGlyPheValGlyGlnAlaAlaAlaArgGluAlaIle 50
|||||
160 AGCGGCGCTGAGGCTGTGTGGGCGAGGAGACCGCGGACAGCATGTGC 209
|||||
51 lValAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaVal 67
:::|
210 GTCATAGTAGAATTAATCAAAACAGAAATGCTGGAAGACCTGCTT 259
|||||
67 uLeuAlaGlyIleProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
|||||
260 GTTGGCAGGACCTCTGGAAGCTGCAAGACAGCTGAGCTGCTGCTAT 309
|||||
84 lArgGlnLeuGlySerLysValProPheCysProMetValGlySerGlu 100
|||||
310 CTCAGAGAGCTGGGTAGTAAGTCCCTTCTGCCCAATGCTGGGAGTGA 359

```



alignment\_block:  
US-09-589-510-4 x HSA010058 ..

Align seg 1/1 to: HSA010058 from: 1 to: 1733

```

1 MetArgIleGluGluValGlnSerThrSerIleuArgIleValTh 17
   ::::::::::::::::::::::::::::::::::::::::::::::
77 ATGAAGATTGAGAGAGTGAAGACACTACGAAGACGCGCATCGCTC 126
   ::::::::::::::::::::::::::::::::::::::::::::::
17 HisThrIleIleuGlyLeuAspAlaAsnGlyMetAlaIleA 34
   ::::::::::::::::::::::::::::::::::::::::::::::
127 CCACAGCCAGCGTGAAGGCTGGGCTGCAGAGCGGCTTGCGCAAGC 176
   ::::::::::::::::::::::::::::::::::::::::::::::
34 LeuAlaAlaGlyPheValGlyGlnAlaAlaArgValAlaAlaGly 50
   ::::::::::::::::::::::::::::::::::::::::::::::
177 AGCGCGCTCAGGCGCTGTGGGCGCCAGAGAAAGCGCGAGGCGATGCGC 226
   ::::::::::::::::::::::::::::::::::::::::::::::
51 LeuAlaValAspMetIleArgGlnIleuIleuMetAlaGlyArgAlaValIle 67
   ::::::::::::::::::::::::::::::::::::::::::::::
227 GTCATAGTAGAATTAATCAAAAGCAAGAAATGGCTGGAAGAGCTGTCTT 276
   ::::::::::::::::::::::::::::::::::::::::::::::
67 uLeuAlaGlyProProlaThrGlyIleThrAlaLeuAlaLeuGlyIleA 84
   ::::::::::::::::::::::::::::::::::::::::::::::
277 GTTGCAGAGACCTCTGGAACCTGCACAGACAGCTCTGCTGCTATTTG 326
   ::::::::::::::::::::::::::::::::::::::::::::::
84 IaGlnGluLeuGlySerIleValProPheCysPrometValGlySerIle 100
   ::::::::::::::::::::::::::::::::::::::::::::::
327 CTCAGGACCTGGGTAGTAGGTCCCTCTGCGCCATGTGGGGAGTGA 376
   ::::::::::::::::::::::::::::::::::::::::::::::
101 ValIleuSerSerGluValIleuIleuThrGluValIleuMetGluAsnPheAr 117
   ::::::::::::::::::::::::::::::::::::::::::::::
377 GTTTCCTCACTAGATCAAGACAGACAGAGTGCATGAGCAACTTCG 426
   ::::::::::::::::::::::::::::::::::::::::::::::
117 GARGAlaIleGlyLeuArgIleuGlyLeuAsnGlyValIleuGlyGly 134
   ::::::::::::::::::::::::::::::::::::::::::::::
427 CAGGCGCATTTGGCTGCATTAAGAGAGCAAGAGATTTATGAAGGTG 476
   ::::::::::::::::::::::::::::::::::::::::::::::
134 JuValThrGluLeuSerProGluIleuIleuIleuThrThrGlyGlyIle 150
   ::::::::::::::::::::::::::::::::::::::::::::::
477 AAGTACAGAGCTAACTCCGTGTGAGACAGAAATCCCATGGGAGATAT 526
   ::::::::::::::::::::::::::::::::::::::::::::::
151 AlaIleuSerIleSerIleValIleIleSerIleuIleuThrValIleuGlyIle 167
   ::::::::::::::::::::::::::::::::::::::::::::::
527 GCGAAACCATTAAGCATGTGATCTAGACTCAAAACAGCCAAAGAGAC 576
   ::::::::::::::::::::::::::::::::::::::::::::::
167 rIleuGlnIleuIleuAspSerSerIleThrAspAlaLeuIleuGlyIle 184
   ::::::::::::::::::::::::::::::::::::::::::::::
577 CAACAGATTGAACCTGACCCGACCATTTTGAAGTTTGCAGAAAGAGC 626
   ::::::::::::::::::::::::::::::::::::::::::::::
184 yValAlaValAlaGlySerValIleuIleuIleuIleuAsnSerGlyValAla 200
   ::::::::::::::::::::::::::::::::::::::::::::::
627 GAGTGAAGAGCTGAGATGTGATTACATTGAGCCACAGAGGCGCGCTG 676
   ::::::::::::::::::::::::::::::::::::::::::::::
201 IysAlaValAlaGlyArgCysAspSerPheAlaThrGluIleuThrAspLeuGluAl 217
   ::::::::::::::::::::::::::::::::::::::::::::::
677 AAGAGGAGGAGGAGGAGTGTGATACCTATGCGCACAGATTCGACCTTGAAGC 726
   ::::::::::::::::::::::::::::::::::::::::::::::
217 aGluGluIleuValProIleuProIleuGlyGluValHisIleuIleuGlyIle 234
   ::::::::::::::::::::::::::::::::::::::::::::::
727 TGAAGAGTATGTCCCTTGCACAAAGGAGATGTGCACAAAGAAAGAA 776
   ::::::::::::::::::::::::::::::::::::::::::::::
234 IeValAlaValAlaThrIleuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
   ::::::::::::::::::::::::::::::::::::::::::::::
777 TCATCCAAAGATGTGACTTGTGCACTTGATGATGGCTAAATGCGGCGCC 826
   ::::::::::::::::::::::::::::::::::::::::::::::
251 GlnGlyGlyGlnAspIleuLeuSerIleuMetGlyIleuMetIleuIleuProAr 267
   ::::::::::::::::::::::::::::::::::::::::::::::
827 CAGGGGGGAGCAAGATATCTGTCCATGATGGCCAGCAATAGAACCCAA 876
   ::::::::::::::::::::::::::::::::::::::::::::::
267 gIleuThrGluIleuThrGluIleuIleuGlnGluIleuAsnIleuValAla 284
   ::::::::::::::::::::::::::::::::::::::::::::::
877 GAAAGCAAGAAATCAAGACAAACTCGAGGAGATTTAATGAAGTGTGTA 926

```

```

284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValIleuPhe 300
   ::::::::::::::::::::::::::::::::::::::::::::::
927 ACAAGTACATCGACACGAGGCAATTGCTGAGCTGCTCCCGGGTGTGCTTT 976
   ::::::::::::::::::::::::::::::::::::::::::::::
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrIleuAsnAr 317
   ::::::::::::::::::::::::::::::::::::::::::::::
977 GTTGATGAGGTCCACATGCTGAGCATTTGAGTGTCTTCACTTACCTGACCG 1026
   ::::::::::::::::::::::::::::::::::::::::::::::
317 gAlaLeuGluSerProLeuSerProIleValIleuAlaThrAsnArgC 334
   ::::::::::::::::::::::::::::::::::::::::::::::
1027 CGCCCTGAGCTTCTTCAATCGCTCCCATGCTGATCTTGATCCACACGAG 1076
   ::::::::::::::::::::::::::::::::::::::::::::::
334 IyIleCysAsnValArgGlyThr...AspMetThrSerProHisGlyIle 349
   ::::::::::::::::::::::::::::::::::::::::::::::
1077 GCACCTGTGTCATCAGAGGACCTGAGGACATCATTCCTCCACGCGCATC 1126
   ::::::::::::::::::::::::::::::::::::::::::::::
350 ProValAspLeuAspArgLeuValIleIleArgThrGluThrIleArgI 366
   ::::::::::::::::::::::::::::::::::::::::::::::
1127 CCTCTGACCTTCTGAGACGAGTGAATGATATCCGACCATGCTGATATC 1176
   ::::::::::::::::::::::::::::::::::::::::::::::
366 yProThrGluMetIleGlnIleuAlaIleArgAlaGlnValGluGlu 383
   ::::::::::::::::::::::::::::::::::::::::::::::
1177 TCCACAGAAATGAACAGATCATTAATAATCCGTGCCACAGCGAAGGAA 1226
   ::::::::::::::::::::::::::::::::::::::::::::::
383 IeAspMetAspGluGluSerIleuAlaThrIleuGlyGluIleuGlyGln 399
   ::::::::::::::::::::::::::::::::::::::::::::::
1227 TCACATCAGTGAAGAGGACCTGACACCTCGGGGAGATGGCACCAAG 1276
   ::::::::::::::::::::::::::::::::::::::::::::::
400 ThrSerLeuArgHisAlaIleGlnIleuIleuSerProAlaSerValIle 416
   ::::::::::::::::::::::::::::::::::::::::::::::
1277 ACCACATGAGGTATCATGATGAGCTGACCTGACCCCGCAACTTGTTC 1326
   ::::::::::::::::::::::::::::::::::::::::::::::
416 rIleuThrAsnGlyArgGluIleuIleuIleuGlyAlaAspLeuGluVal 433
   ::::::::::::::::::::::::::::::::::::::::::::::
1327 TAAATACACGGGAGGAGACAGATTTGAAGAAAGCATGTGCAAGATCA 1376
   ::::::::::::::::::::::::::::::::::::::::::::::
433 ergIleuIleuThrLeuAspAlaIleuSerSerAlaArgLeuGluGln 449
   ::::::::::::::::::::::::::::::::::::::::::::::
1377 GTGACCTTTCTATGATGCCAAGTCTCTCGCCAAATCTCGGTGACAG 1426
   ::::::::::::::::::::::::::::::::::::::::::::::
450 GlnGluArgTyrIle 454
   ::::::::::::::::::::::::::::::::::::::::::::::
1427 CAGGATTAAGTACATG 1441

```

seq\_name: gb\_p14:AF099084

seq\_documentation\_block:

LOCUS AF099084 1736 bp mRNA PRI 13-JAN-1999

DEFINITION Homo sapiens Pontin5f mRNA, complete cds.

ACCESSION AF099084

VERSION AF099084.1 GI:4151524

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1736) Bauer, A., Huber, O., and Kemler, R.

Pontin52, an interaction partner of beta-catenin, binds to the TATA

box binding protein Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14787-14792 (1998)

JOURNAL 99061945

REFERENCE 2 (bases 1 to 1736) Bauer, A., Huber, O., and Kemler, R.

Direct Submision

Submitted (19-OCT-1998) Department of Molecular Embryology,

Max-Planck Institute of Immunobiology, Stuebeweg 51, Freiburg

D-79108, Germany

FEATURES location/Qualifiers

SOURCE

1..1736 /organism="Homo sapiens"

/db\_xref="taxon:9606"



```

/Cell_line="SM480"
/tissue_type="colorectal adenocarcinoma"
65..1435
/function="binds to the TATA-box-binding protein (TBP)"
/codon_start=1
/product="Ponlin52"
/db_xref="GI:4151525"
/translation="MKIEVKSTKTORTASHHYKGLIDESGLAKQASGLVGOEN
AREACGIVIELLSKKMAGRAVLLAGPRTGTALALAIQELSGKVPFCMVSEVY
STEIKTEVLMENFRRAIGLRKETKEVEYEGEVLELTPCETENPKGIGKTSVHII
LKTAGTQKQLDPSIFESLQKREVEADVIYEANSVNRQGRCDYATPEFLAE
EXVPLPKGVHKKKEIIDVTLHDVDANARPOGODILSMGQLMKRKTEIDKLK
GEINKVYMKYIDOGIAELVPGVLEFVEYEMDIEFTLHRLAESLAPVIYFASNRG
NCVINGTEDITSPHGIDPLDRLVMTITRMLTPOEMKOITKIRAOFTGISEALN
HGEIGKTTLRISYQVLLTPANLLAKINGKDSIEKEHVEITSELYDAKSSKILLADQ
ODTKMK"
BASE COUNT      473 a      400 c      481 g      382 t
ORIGIN
alignment_scores:
    Quality: 1723.50      Length: 455
    Ratio: 4.113      Gaps: 1
Percent Similarity: 92.088      Percent Identity: 72.527
alignment_block:
US-09-589-510-4 x AF099084 ..
Align seg 1/1 to: AF099084 from: 1 to: 1736
1 MetArgIleGluGluValAlaGlnSerThrSerLysGlnArgIleAlaThr 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 ATGAGATTTAGAGAGGTGAAGAGCACTACAGACAGCGCATCGCCTC 114
17 RHSTHHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
115 CCACAGCGCAGCTGAAGGCGCTGGCGGAGAGCGGCTTGCCCAACG 164
34 LalaValAlaGlyPheValGlyLalaAlaAlaAlaArgIleAlaGly 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
165 AGCGGCGCTCAGGCGCTTGCGCCAGAGAACGCCGAGAGCGCATGTGGC 214
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyValAlaVal 67
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 GTCATATGATTAATTAACAAGACAGAAATGCTGGAAGAGACTGTCTT 264
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIle 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 GTTGGCGAGGACTCTCGGAACGTGGCAAGACAGCTGCTGCTGCTATTG 314
84 LalaGlnLysLysLysLysValProPheCysProMetValGlySerGln 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
315 CTCAGAGGCTGGGTAGTAAGTCCCTCTGCCCCAATGGTGGAGGTGA 364
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
365 GTTTACCACTGAGATCAAGAAGACAGAGCTCTCATGAGAGACTTCGG 414
117 gArGAlaIleGlyLeuArgIleLysGluAsnGlyValTyrGluGlyG 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
415 CAGGCGCATTTGGCTCGAATAAAGAGACCAAGGAAGTTTATGAAGTG 464
134 LuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyGly 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
465 AAGTCAAGAGGCTAAGCTCGGTGAGACAGAGATCCATGGAGGATAT 514
151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
515 GGCAGAAACCTTAGCATGTGATGAGACTCAAAACAGCCAAAGAAC 564
167 rLysGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

565 CAACAGTTGAACCTGACCCCGACATTTTGAAGTTTCAGAAAGAGC 614
184 ysValAlaValAlaGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
615 GAGTAAAGCTGAGAGATGTGATTTACATTGAAGCCAAACAGTGGCCGTG 664
201 LysArValAlaGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
665 AAGAGCGCAGGCGCAGGTGTATACCTATGCGCACAGAAATTCAGACTTGAAGC 714
217 aGluGluTyrValProIleProLysGlyGluValHisLysLysGluI 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
715 TGAAGAGTATGTGCCCTTCCAAAAGGGAGATGCACAAAAAGAAAGAAA 764
234 lValAlaGlnAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
765 TCATCCAGATGTGACCTTGACATGCTTGATGTGGTAATGCGGGGCC 814
251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
815 CAGGCGGAGCAAGATATCTCTGTCATGATGGCCAGCTAATGAGCCAAA 864
267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValAla 284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
865 GAAGACAGAAATCACAGCAAACTTCAGAGGAGATTAATAAGGTGTGA 914
284 snArgTyrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
915 ACAAGTACATGACCAAGGCGCATTTGCTGAGCTGTCCGGGTGTGTTT 964
301 IleAspGluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
965 GTTGATGAGATGTCACATGCTGAGCATTTGAGTGTCTTACCTACCTGCACAG 1014
317 gAlaLeuGlnSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1015 CCCCCGTGGAGTCTTATACGCTCCCATGTCATCTTTCATCCAAACCGAG 1064
334 lYleCysAsnValAlaArgLysThr...AspMetThrSerProHisGlyIle 349
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1065 GCACTGTGTCTACAGAGCGACTGAGAGCATACATCCCTCAGCGGCAATC 1114
350 ProValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrG 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1115 CCTCTGACCTCTGACCGAGATGATGATATCCGAGACATGCTGATATAC 1164
366 yProThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGlu 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1165 TCCACAGAGAAATGAACAGATCATTAATAATCCGTCGACGAGAGAGAA 1214
383 leAspMetAspGluGlnSerLeuAlaTyrLeuGlyGluIleGlyGlnGln 399
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1215 TCACATCATGAGGAGGAGGAGCAACACCTGGGGGAGATTGGACCCAG 1264
400 ThrSerLeuArgHisAlaIleGlnLeuIleSerProAlaSerValAla 416
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1265 ACCACACAGAGGTACTCATGCTCAGCTGCTGACCCCGCAACTTGCTTGC 1314
416 rLysThrAsnGlyArgGluLysIleCysLysAlaAspLeuGlnGluValS 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1315 TAAATCAACGCGGAAGGACACATGTAGAAAGACATGTCGAAAGAGATCA 1364
433 eArgLysLeuAspAlaLysSerSerAlaArgLeuLeuGlnGlnGln 449
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1365 GTGAAGCTTTTATGATGCAAGTCTCTCCGCAAAATCTGTGACACAG 1414
450 GlnGluArgTyrIle 454
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1415 CAGGATAGTACATG 1429
seq_name: gb_p7.BC002993
seq_documentation_block:

```



LOCUS BC002993 1746 bp mRNA PRI 16-MAR-2001  
 DEFINITION Homo sapiens, Ruvb (E coli homolog)-like 1, clone MGC:3762, mRNA, complete cds.  
 ACCESSION BC002993  
 VERSION BC002993.1 GI:12804268  
 KEYWORDS MGC.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1746)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 CONTACT: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: [Robert.Strausberg@nih.gov](mailto:Robert.Strausberg@nih.gov)  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada  
 info@bgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guln, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRML Plate: 9 Row: K Column: 14.  
 Location/Qualifiers  
 1. 1746  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:3762 IMAGE:2823568"  
 /tissue\_type="lung, small cell carcinoma"  
 /clone\_11b="NH-MGC\_7"  
 /lab\_host="DH10B-R"  
 /note="Vector: pORF7"  
 72. 1442  
 /codon\_start=1  
 /product="Ruvb (E coli homolog)-like 1"  
 /protein\_id="AAH02993.1"  
 /db\_xref="GI:12804269"  
 /translation="MKIEVKSTKTORTLASHSHVKGILDESGILAKOASGLVGOEN AEEAGVIVELIKSKRMAGRAVLGAGPCTGATLALAIAGELSKVPCMCSEVY STEIKTEVLMENFRATGLRIKETREYEGVETLTTCETENPBGCTGKTTSHYITG EYTPKGTQKLDPSIFESLQKEREADVITENSAAVNRQGGCDTGAETFEDEAE LKVRPLKGDVHKKEIIOVDVLHLDVANNAROGQDILSMGQIMKPKKEITDKLR GEINKVNVKIDOGIAELVPGVLFVDEVMHLDIECTYLIHRLAESSTIPIYFASNRG NCVRGTEDTSPHGIPLDLDRVMIIRMTLYTPEMKOIRKIRQOTGGINISEALN HGEIGTKTTLTLYSVQLTPANLLAKINKGDSIEKHEVEISELFYDAKSSAKILLADO QPKYK"  
 BASE COUNT 479 a 403 c 479 g 385 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 1723.50 Length: 455  
 Ratio: 4.113 Gaps: 1  
 Percent Similarity: 92.088 Percent Identity: 72.527  
 alignment\_block:

US-09-589-510-4 x BC002993 ..  
 Align seg 1/1 to: BC002993 from: 1 to: 1746  
 1 MetatrgIleGluGluValGlnSerThrSerLysLysGlnArgIleAlaIath 17  
 72 ATGACGATTGAGGAGGTGAGACACTACGACAGACGACGACGACGCTC 121  
 17 rHsThrIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34  
 122 CCACAGCCACGTGAAGGGCTGGGCTGGAGCAGAGCGGCTTGGCCAAAGC 171  
 34 LeuAlaAlaGlyPheValGlyGlnAlaAlaAlaArgGluAlaGly 50  
 172 AGCGGCGCTCAGGCGCTTGGGCGCAGAGAACGCGGAGAGGATGTGGC 221  
 51 LeuAlaAlaAspMetIleArgGlnLysLysMetAlaGlyArgAlaValle 67  
 222 GTCATGTGTAATTAATCAAAAGCAAAAGAAATGGCTGGAAGAGCTGCTT 271  
 67 IuEuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84  
 272 GTTGGCAGGACCTCTGGAAGCTGGACAGACAGCTCTGGCTGCTATTG 321  
 84 IAGlnGluLeuGlySerLysValProPheCysProMetValGlySerGlu 100  
 322 CTCAGGAGCTGGGTAGTAGTCCCTTCTGCCAATGGTGGGAGGTGA 371  
 101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAspPheAr 117  
 372 GTTACTCACTGACATCAAGAACAGAGAGGTGCTGATGAGAACCTTCG 421  
 117 grrgAlaIleGlyLeuArgIleLysGlnLysGluValTyrGlyGlyG 134  
 422 CAGGGGCAATTGGCTGCGAATTAAGGAGACCAAGAGATTATGAGAGGTG 471  
 134 IuValThrGluLeuSerProGluGluAlaGluSerThrThrGlyLysTyr 150  
 472 AAGTCACAGACGTAACCTCGGTGGAGACAGACATCCATGGGAGATAT 521  
 151 AlaLysSerIleSerHisValIleIleSerLeuLysThrValLysGlyTyr 167  
 522 GGCAAACCATTAAGCATGTGATCATAGACATCAAAACAGCCAAAGAAC 571  
 167 rlyGlnLysLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184  
 572 CAACAGTTGAAACCTGACCCGACCATTTTGAAGTTTGACAGAAAGAC 621  
 184 ysvAlaAlaValGlyAspValIleTyrIleGluAlaAsnSerGlyAlaVal 200  
 622 GAGTGAAGCTGGAGATGTGATTTACATTGAAGCCACAGTGGGCGCTG 671  
 201 LysArgValGlyArgCysAspSerPheAlaThrGluTyrAspLeuGluAl 217  
 672 AAGAGCAGGACGAGGTGTGATACCTATGACCAAGAAATTTGACCTTGAAGC 721  
 217 agLgLuTyrValProIleProLysGlyLysValHisLysLysGlu 234  
 722 TGAAGGAGTGTGCCCTTCCCAAAAGGGATGTGCACAAAGAAAGAAA 771  
 234 IeValAlaAspValThrLeuHisAspLeuAspAlaAlaAsnAlaGlnPro 250  
 772 TCATCCAAAGTGTGACCTTGATGATTTGATGTGGCTAATCCGCGGCC 821  
 251 GlnGlyGlyGlnAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267  
 822 CAGGGGAGCAAGATATCTGTCATGATGAGGCGCAGCTAATGAAGCAAA 871  
 267 gLysThrGluIleThrGluLysLeuArgGlnGluIleAsnLysValValA 284  
 872 GAAGACAGAAATTCACAGACAACTTGGAGGAGGATTAATTAAGGTGGTA 921  
 284 snArgTyrIleAspGluGlyIleAlaGluValProGlyValLeuPhe 300

```

|||||
922 ACAAGTACATGACAGGCGATTGCTGAGCTGGTCCGGGTGCTGTTT 971
301 l leaspluValHisMetLeuAspIleGluCysPheSerTyrLeuAsnAr 317
972 GTTGATAGGTCCACATGCTGACATTTGATGCTTACCTACCTGACCG 1021
317 gAlaLeuGluSerProLeuSerProIleValIleLeuAlaThrAsnArg 334
1022 CGCCCTGGAGCTTTCATCGCTCCCATCGTCATCTTGTGATCCACCGAG 1071
334 l ylleCysAsnValArgGlyThr... AspMetThrSerProHisGlyIle 349
1072 GCACACTGTGATCAGAGGACAGGACATGACATCCCTCAGGCGATC 1121
350 ProValAspLeuLeuAspArgLeuValIleIleArgThrGluThrTyrG 366
1122 CCTTGTGACCTCTGGACCGAGTATGATATCCGACCATGCTGTATAC 1171
366 yProThrGluMetIleGlnIleLeuAlaIleArgAlaGlnValGluGlu 383
1172 TCCACAGGAATGAAACAGATCATTAATCCGTGCCAGAGAGAGAA 1221
383 leasPheLaspGluSerLeuLeuAlaTyrLeuGlyGluIleGlnGln 399
1222 TCACATCATGAGGAGGACAGTACACCTGGGGAGATGGCACCAAG 1271
400 ThrSerLeuArgHisAlaIleGlnLeuIleSerProAlaSerAlaValSe 416
1272 ACCACACTGAGTACTCAGTGCAGCTGTCGACCCCGCCACACTGCTTGC 1321
416 rlyStrAsnGlyArgGluIleGlySerLysAlaAspLeuGluValS 433
1322 TAAATTAACGCGGAGACAGACATGTGAAAGCATGTGCAAGAGATCA 1371
433 excIleuTyrLeuAspAlaLysSerSerAlaArgLeuLeuGlnGln 449
1372 GTAACTTTTCTATGATGCCAGTCTCCGCCCAAAATCTGTGCTGACAG 1421
450 GlnGluArgTyrIle 454
1422 CAGGATAGTACATG 1436

seq_name: gb_pr4:AF070735
seq_documentation_block:
LOCUS AF070735 1750 bp mRNA PRI 17-NOV-1998
DEFINITION Homo sapiens RuvB-like protein RUVBL1 mRNA, complete cds.
ACCESSION AF070735
VERSION AF070735.1 GI:3243034
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1750)
Qiu,X.-B., Lin,Y.-L., Thome,K.C., Pian,P., Schlegel,B.P.,
Weremowicz,S., Parvin,J.D. and Dutta,A.
An eukaryotic RuvB-like protein (RUVBL1) essential for growth
J. Biol. Chem. 273 (43), 27786-27793 (1998)
98447618
MEDLINE
REFERENCE
2 (bases 1 to 1750)
Qiu,X.-B., Lin,Y.-L., Thome,K.C., Pian,P., Schlegel,B.P.,
Weremowicz,S., Parvin,J.D. and Dutta,A.
Direct Submission
JOURNAL
Submitted (05-JUN-1998) Pathology, Brigham & Women's Hospital,
Harvard Medical School, 75 Francis Street, Thon 630, Boston, MA
02115, USA
FEATURES
source
1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"

```

```

CDS
/mmap="3q21"
77..1447
/codon_start=1
/product="RuvB-like protein RUVBL1"
/protein_id="AAC7819.1"
/db_xref="GI:3243035"
/translaton="MKIEEVKSTTKTORIASHSVHVGILDESGLAQQAASGLVGOEN
ARACGVIVLEILKSKMAGRAVLVLAGPFGGTALALAIQELGSKVPEPCPMVGSEVY
STRIKTEVIMEMFRRARIGRIKETKEVEGEVTELTPECTENPMGCKYKTSYITIG
LTKAKSTKQIKLDPSTFESIQKRRVAGVITYEANSQAVKRGKRDYATFELEAE
EYVLEKRGDVHKKKEIIODYTLHLDLVANAARQGGDILSMGQLMKPKRTITOKLR
GELINKVMRYIDGGLAEVLVEDEVMHLDIECTYDLRALLESSIAPIVFAANRG
NCVIRGTEITDSHGIPDLDRVMIIRFMYTPQEMKQIIKIRATQESINSEBAL
HGEISTKTTIYRSVQLLPANLAKINKGDSIEKHVEISELPEADKSSAKIILADQ
ODKYMK"

BASE COUNT      476 a      403 c      484 g      387 t
ORIGIN

alignment_scores:
Quality: 1723.50      Length: 455
Ratio: 4.113      Gaps: 1
Percent Similarity: 92.088      Percent Identity: 72.527

alignment_block:
US-09-589-510-4 x AF070735

Align seg 1/1 to: AF070735 from: 1 to: 1750

1 MetArgIleGluGluValGlnSerThrSerLysLysGlnArgIleAlaThr 17
|||||
77 ATGAAGATTGAGAGAGTGGAAGAGCATACGAAGACGACGCGCATCGCTTC 126
|||||
17 rHisThrHisIleLysGlyLeuGlyLeuAspAlaAsnGlyMetAlaIleA 34
|||||
127 CCACAGCCAGTGAAGAGGCGTGGGCTGGACGAGAGCGCTTGGCCAAAGC 176
|||||
34 lAlaAlaAlaGlyPheValGlyGlnIleAlaAlaArgGlnAlaAlaGly 50
|||||
177 AGGCGGCTCTCAGGCGCTTGTGGCCAGAGAGAACGCGCAGAGCATGTGGC 226
|||||
51 LeuAlaValAspMetIleArgGlnLysLysMetAlaGlyArgAlaValle 67
|||||
227 GTCATAGTAGAATTAATCAAAAGCAAGAAATAGCTGGAAGACCTGTCTT 276
|||||
67 uLeuAlaGlyProProAlaThrGlyLysThrAlaLeuAlaLeuGlyIleA 84
|||||
277 GTTGGCAGAGACTCTCTGGAAGCTGGAAGACAGCTGTGCTGTGGCTATTTG 326
|||||
84 lArgIleLeuGlySerLysValProPheCysProMetValGlySerGlu 100
|||||
327 CTCAGGAGCTGGGTAGTAAGGTCTCCCTTCTGCCCAATAGTGGGAGATGAA 376
|||||
101 ValTyrSerSerGluValLysLysThrGluValLeuMetGluAsnPheAr 117
|||||
377 GTTACTCAACTGAGATCAAGAAGACAGAGAGTCTCATGAGAACTTCGC 426
|||||
117 gArgAlaIleGlyLeuArgIleLysGlnAsnLysGlnValTyrGlnGlyG 134
|||||
427 CAGGCGCATTTGGGCTGGCAATTAAGGACCAAGGAAGTTTAAGAGGTG 476
|||||
134 l uValThrGluLeuSerProGluAlaGluSerThrThrArgTyrIle 150
|||||
477 AAGTCACAGAGCTAAGTCCGTGTGAGACAGAGATATCCATGGAGATAT 526
|||||
151 AlalysSerLieserHisValIleIleSerLeuLysThrValLysGlyTh 167
|||||
527 GCCAAACCATTAAGCATGATGATAGACCTCAAAACAGCCAAAGAAC 576
|||||
167 rlyGlnLeuLysLeuAspSerSerIleTyrAspAlaLeuIleLysGlu 184
|||||
577 CAACAGTTGAACCTGACCCCGACATTTTGAAGTTCGAGAAAGAGC 626

```

ACCESSION	Y18418
VERSION	Y18418.1
KEYWORDS	ECP-54; erythrocyte cytosolic protein.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1789)
AUTHORS	Salzer,U., Kubicek,M. and Prohaska,R.
TITLE	Isolation, molecular characterization, and tissue-specific expression of ECP-51 and ECP-54 (TIP49), two homologous, interacting erythroid cytosolic proteins
JOURNAL	Biochim. Biophys. Acta 1446 (3), 365-370 (1993)
MEDLINE	99453736
REFERENCE	2 (bases 1 to 1789)
AUTHORS	Salzer,U.
TITLE	Direct Submission
JOURNAL	Submitted (13-NOV-1998) U. Salzer, Institute of Biochemistry, University of Vienna, Vienna Biocenter, Dr. Bohr-Gasse 9/3, A-1030 Vienna, AUSTRIA
FEATURES	Location/Qualifiers
source	1..1789
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/tissue_type="bone marrow"
	<1..96
5'UTR	97..1467
CDS	/codon_start=1
	/product="erythrocyte cytosolic protein of 54 kDa, ECP-54"
	/protein_id="CAB46271.1"
	/db_xref="GI:5327000"
	/translation="MKIEEYKSTTKTORIAASHNVKIGLDESGILAKQAASGLNGEN ARACQIVELIKSKMAGRAVILLACPRTGSHALALALADELGSVPFCPMWGSFV SEIRKTDEVIMENFRRAIGRIKETEVTEGETELTPETENPWSGVGTTSHTVIG LKAKRGKQLDPSIFESIDOKERVAEGOVITYEANGSAVKRGRCDVTATFEDEAE EYPLRKGGVDHKREKLEIIOVTLADLDVANARPQGODLISMMGOILMKPKKTEITTKLR GEINKVNNKYIIDGIELAELPVGLVEDEVMLEICTPTYLHALESSIAPVTFASNKG NCSIQREDITSHPGIRPLDLDRVMILRTMLVTPQPMKOITIKRAOTEGINISEALN HLEICETKTTLRSVOLLPANILANKDKSIDEKHVHEISBELPYDAKSSAKILIADQ ODVMK"
3'UTR	1468..>1789
polyA_signal	1737..1742
BASE COUNT	496 a 412 c 491 g 390 t
ORIGIN	
alignment_scores:	
Quality:	1723.50 Length: 455
Ratio:	4.113 Gaps: 1
Percent Similarity:	92.088 Percent Identity: 72.527
alignment_block:	
US-09-589-510-4 x HOSAI8418 ..	
Align seg 1/1	to: HOSA18418 from: 1 to: 1789
1	MeLArgllEglugluValGlnSerThSerTyslsgInArgrlaleatrh 17
	ATTGAAGAATTTGAGGAGTGCAAGACATCAAGACGCACATCGCCCTC 146
97	ATGGAACAATTTGAGGAGTGCAAGACATCAAGACGCACATCGCCCTC 146
17	rHisThrhisIllelysgIleungIytleuAspAlaAsnGlMeFAllaLea 34
	:::::     :::::     :::::     :::::     :::::      196
147	CCAACAACCACGTGAAGAGCGCTGTGGGCTGGAGACAGCGCGCTTGCCCAAGC 196
34	IalEuAlalaIagLyPheValGlyslnAlaAlaAlaArgIuAlaIagLy 50
	:::::     :::::     :::::     :::::     ::::: 197
197	AGGCGGCGCTCAGGCGCTTGTGGGCCAGCAAGAACGCGCAGAGCATGTGGC 246
51	leuAlaIalaspmEtIIeArGlnIyIsgMeFAIargIlaIaValle 67
	:::      :::::     :::::     :::::     :::::      247
247	GTCATAGTAGTAATTAAITCAAAGCAAGAAAATGGCTGGAAAGACTGTCTTT 296
67	uLeuAlaIagLyProFrAlathrGlyIystrAlaIeuaAlaIeunGlYlea 84

```

|||||
297 GTTGGCAGACCTCTGCACTGGCAAGACAGCTCTGCTGCTATG 346
84 ladingluenclgyserysvalprophcysprometvalgl1serglu 100
|||||
347 CTCAGAGCTGGGTAGTAGTCCCTTCTGCTCCATGTGGGGAGTGA 396
397 GTTACTCACTGAGATCAGAGACAGAGAGTGTGATGAGAACTTCCG 446
101 valtyrsersergluvallyslysThrGluValLeuMetGluAsnphar 117
|||||
117 gATgAlalleglyleuAryllelysgluAsnlysgluValtyrGlu 134
447 CAGGGCCATTGGCTGCGAATAAGAGACCAAGAGATTATGAGGTG 496
134 luValThrcGluLeuSerProGluGluValGluSerThrGlyGlyTyr 150
|||||
497 AAGTCAAGAGCTAACTCCGTGTGAGACAGAGAAATCCATGGGAGATAT 546
151 AlAlysSerThrSerHisValIleIleSerLeuLysThrValLysGlyTh 167
|||||
547 GCAAAACCATTAAGCATGATGATCATAGACTCAAAACAGCCAAAGAGAC 596
167 rlysgluLeuLysLeuAspSerSerIleTyraAspAlaLeuIleLysGlu 184
597 CAACAGGTGAACTGAGACCCAGCATTTTAAAGTTTGCAGAAAGAGC 646
184 ysValAlaValAlglYasPvalIleTyrllegluAlaAsnSerGlyAlaVal 200
647 GAGTAAAGCTGAGATGTGATTACATTGAGCAACAGTGGGCGCGTG 696
201 LysAlglValAlglYarGysAspSerPheAlaThrGluTyraAspLeuGluAl 217
697 AAGAGCCAGGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 746
217 agluGluTyraValProIleProLysGlyGluValHisLysLysGlyLut 234
747 TGAAGATGTATGCTCCCTTCCAAAGGGATGTGCACAAAGAAAGAAAGAA 796
234 lvalAlglAspValThrLeuHisAspLeuAspAlaAlaAsnLacIlePro 250
797 TCATCCAGATGTGACCTTGATGATGATGATGATGATGATGATGATGATGATG 846
251 glnglyGlyGluAspIleLeuSerLeuMetGlyGlnMetMetLysProAr 267
847 CAGGGGGGAGAAAGATATCTGTCATGATGAGGCGCACTATATAGCCANA 896
267 glysThrGluIleThrGluLysLeuArgGlnGluLeuValLeuVal 284
897 GAAGACAGAAATCAGACAGAACTTGAGGGGAGATTATATAGTGTGA 946
284 snArGlyThrIleAspGluGlyIleAlaGluLeuValProGlyValLeuPhe 300
947 ACAAGTACATCGACAGGCGATGCTGAGCTGCTCCGGGTGTGCTGT 996
301 lIasGluValHisMetLeuAspIleGluLysPheSerTyrlLeuAsnAr 317
997 GTGATGAGGTCCACATGCTGACATGAGTGTGATGATGATGATGATGATGATG 1046
317 gAlaLeuGluSerProLeuSerProIleValIleLeuValThrAsnArgG 334
1047 CGGCTGAGATCTTTCATGCTCCATGCTCCATGCTTTCATGCTCCATGCTCC 1096
334 lylleCysAsnValArgGlyThr..AspMetThrSerProHisGlyIle 349
1097 GCAACTGTGTCTATCAGAGGCGATGAGAGATCATCCCTCAGGCGATC 1146
350 ProValAspLeuLeuAspArgLeuValIleLeargThrGluThrTyrgl 366
1147 CCTCTTACCTTCTGAGCGAGTATGATATATCCGACCATGCTGTATATAC 1196
366 yProThrGluMetIleGlnIleLeuValAlaArgAlaGlnValGluGlu 383
|||||

```

```

1197 TCACAGAGAAATGAAACAGATCATTAATAATCCGTGCCAGAGGAGAA 1246
383 leasphetaspgluGluSerLeuAlaIleTyrlleuGlyLutIleGlyGln 399
1247 TCACATCATGTGAGAGGAGCTGACACCTGGGAGATGAGACCAAG 1296
400 ThrSerLeuArgHisAlaIleGlnIleLeuIleSerProAlaSerValAla 416
1297 ACCACACTGAGTACTGACGACGCTGACCCCGGCCAACATGCTTGC 1346
416 rlyThrAsnGlyArgGluLysIleCysLysAlaAspLeuGluValS 433
1347 TAAATCAACGGAGAGACAGCATTAAGAAAGCATGTGCAAGATCA 1396
433 ergLeuTyrlleuAspAlaLysSerSerLarArgLeuGlnGluGln 449
1397 GTGACCTTTCTATGATGCCAGTCTCCGCAAAATCTGCTGACCA 1446
450 GlnGluArgTyrlle 454
1447 CAGGATTAAGTACATG 1461
seq_name: gb_pr7:BC000519
seq_documentation_block:
LOCUS BC000519 2503 bp mRNA PRI 16-MAR-2001
DEFINITION Homo sapiens, RUvB (E coli homolog)-like 1, clone MGC:8557, mRNA, complete cds.
ACCESSION BC000519
VERSION BC000519.1 GI:12653494
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
nisc.mgcgenbri.nih.gov
Contact:
Shevchenko,Y., Wehrt,D.K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Snyder,B., Stantropop,S., Thomas,P.J.,
Tjongson,E.E., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A.,
Zhang,L.-H. and Green,E.D.
FEATURES
SOURCE
location/Qualifiers
1..2503 = "Homo sapiens"
/db_xref="locusID:8607"
/db_xref="taxon:9606"
/clone="MGC:8557 IMAGE:2822889"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NIH_MGC-7"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 1 Row: 1 Column: 20.

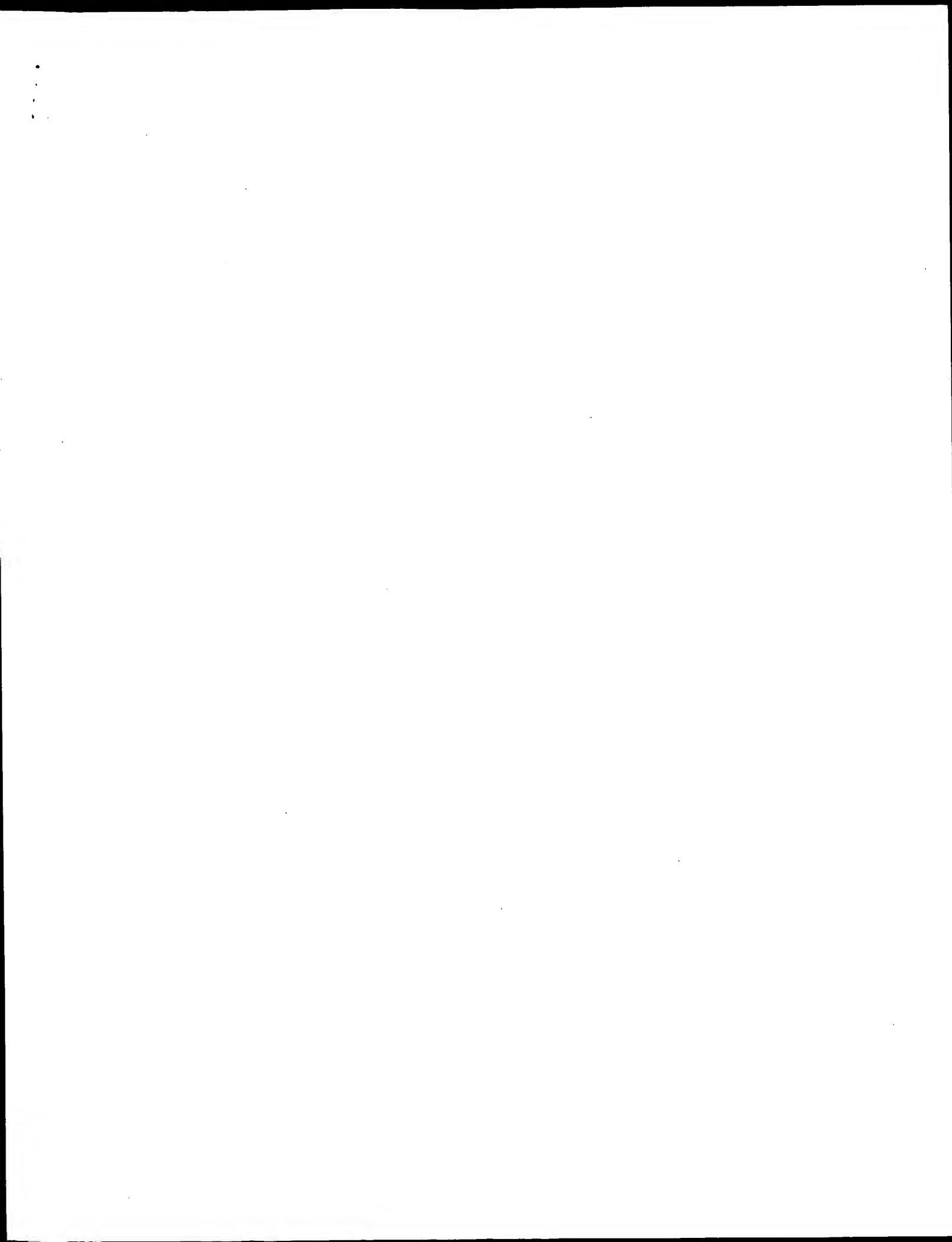




```

1068 CCCTTGACCTGCTGGACCGGGTGCATGATCATCATGACACCATGCTGTATAC 1117
366 YProthrgiUmeIleGlnIleLeuAlaIleArgAlaIleValGluIu 383
1118 GCCACAGGAGATTAACCAACATCATTTAAAGTCCGAGCCAGACGGAAGCA 1167
383 leasPwelaSpGluIuSerLeuAlaTyrLeuGluIleGluGln 399
1168 TCAACATCAGTGGAGGAGGCCCTTAACCAACCTGGGGAATTTGGCACCAAG 1217
400 ThSerLeuArhIAlaIleGlnLeuIleSerProAlaSerValIaIe 416
1218 ACCACGCTGAGTATTCAGTCGACGCTGCAGACCTCGCAACTGTGGC 1267
416 rlySThAsnGluArhGluIuIleCysIyAlaAspLeuGluIuAlS 433
1268 CAGAGATCAACGGGAAGACACGATTCAGAGAGAGACACGTGAGGAGATCA 1317
433 ergIlyLeuTyrLeuAspAlaIySSerAlaArgLeuLeuGlnGln 449
1318 GCGAGCTCTTCATGACGCGCAAGTCTCCGCCAGAGATTCTGGCCGACAG 1367
450 GlnGluArgTyrIle 454
1368 CAGGACAGATCAATG 1382

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

On nucleic - nucleic search, using sw model

Run on: November 13, 2001, 10:42:30 Search time 143.43 Seconds  
(without alignments)  
8370.270 Million cell updates/sec

Title: US-09-589-510-3

```

Perfect score: 1912
Sequence: 1 accacgcgtccgcaatt.....aaaaaaaaaaaaaaaaa 1912

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
1: /SIDSL/gcgdata/geneseq/genseq/NAl1980.DAT *
2: /SIDSL/gcgdata/genseq/genseq/NAl1981.DAT *
3: /SIDSL/gcgdata/genseq/genseq/NAl1982.DAT *
4: /SIDSL/gcgdata/genseq/genseq/NAl1984.DAT *
5: /SIDSL/gcgdata/genseq/genseq/NAl1984.DAT *
6: /SIDSL/gcgdata/genseq/genseq/NAl1985.DAT *
7: /SIDSL/gcgdata/genseq/genseq/NAl1986.DAT *
8: /SIDSL/gcgdata/genseq/genseq/NAl1987.DAT *
9: /SIDSL/gcgdata/genseq/genseq/NAl1988.DAT *
10: /SIDSL/gcgdata/genseq/genseq/NAl1989.DAT *
11: /SIDSL/gcgdata/genseq/genseq/NAl1990.DAT *
12: /SIDSL/gcgdata/genseq/genseq/NAl1991.DAT *
13: /SIDSL/gcgdata/genseq/genseq/NAl1992.DAT *
14: /SIDSL/gcgdata/genseq/genseq/NAl1993.DAT *
15: /SIDSL/gcgdata/genseq/genseq/NAl1994.DAT *
16: /SIDSL/gcgdata/genseq/genseq/NAl1995.DAT *
17: /SIDSL/gcgdata/genseq/genseq/NAl1996.DAT *
18: /SIDSL/gcgdata/genseq/genseq/NAl1997.DAT *
19: /SIDSL/gcgdata/genseq/genseq/NAl1998.DAT *
20: /SIDSL/gcgdata/genseq/genseq/NAl1999.DAT *
21: /SIDSL/gcgdata/genseq/genseq/NAl2000.DAT *
22: /SIDSL/gcgdata/genseq/genseq/NAl2001.DAT *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1910.8	99.9	1912	22	AAD02567	Maize RVB orthoid
2	1686.6	88.2	1845	22	AAD02566	Maize RVB orthoid
3	1630	85.3	1869	22	AAD02570	Maize RVB orthoid
4	1362.2	71.2	1886	22	AAD02568	Maize RVB orthoid
5	1351.8	70.7	1898	22	AAD02569	Maize RVB orthoid
6	783.6	41.0	1612	21	AAC38992	Arabidopsis thailia
7	622.6	32.6	1587	20	AAK18198	TP49 coding sequ
8	612.4	32.0	1730	20	AAK18199	TP49 coding sequ
9	612.4	32.0	1750	20	AAK24059	Human regulatory m
10	544.4	28.5	1744	21	AAC59220	Human secreted prote
11	524	27.4	1509	20	AAZ52915	Human prostate tum

12	36.8	17.6	121.3	21	AA137370	Aspergillus oryzae
13	244.6	12.8	1416	21	AA930317	Yeast HEL50 DNA he
14	228.2	11.9	1492	21	AA930316	Human HEL50 DNA he
15	166	8.7	1280	22	AA874190	DNA encoding eviro
16	157.8	8.3	1167	20	AA874507	Human secreted pro
17	137.2	7.2	335	21	AA808054	Fusarium venenatum
18	129.6	6.8	1002	20	AA930161	Human secreted pro
19	128.4	6.7	429	21	AA802098	Human secreted pro
20	118.8	6.2	433	21	AA804039	Human secreted pro
21	109.6	5.7	749	21	AA807528	Human secreted pro
22	108.6	5.7	313	14	AA060576	Fusarium venenatum
23	101.6	5.3	936	22	AAE58257	Human brain expres
24	101.4	5.3	936	22	AAE58254	Oligonucleotide D1
25	101.4	5.3	936	22	AAE58255	Oligonucleotide D1
26	101.2	5.3	936	22	AAE58252	Oligonucleotide D1
27	99.8	5.2	936	22	AAE58259	Oligonucleotide D2
28	99.8	5.2	936	22	AAE58262	Oligonucleotide D2
29	98.6	5.2	936	22	AAE58252	Oligonucleotide D1
30	98.6	5.2	936	22	AAE58254	Oligonucleotide D1
31	98.6	5.2	936	22	AAE58257	Oligonucleotide D1
32	98.6	5.2	936	22	AAE58259	Oligonucleotide D1
33	98.6	5.2	936	22	AAE58262	Oligonucleotide D2
34	98.6	5.2	938	22	AAE58255	Oligonucleotide D2
35	54.6	2.9	769	21	AA828122	Oligonucleotide D1
36	53	2.8	844	21	AA829112	Arabidopsis thalia
37	51.4	2.7	721	21	AA829112	cDNA encoding C-te
38	51.4	2.7	1696	21	AAE16214	Human secreted pro
39	51.4	2.7	3891	21	AAE74624	Human prostate can
40	51.2	2.7	309	21	AA894021	Human ORF1 ORF1979
41	51.2	2.7	310	21	AA894563	Cat flea hindgut a
42	50.4	2.6	8749	22	AA887129	Cat flea hindgut a
43	50.4	2.6	8941	22	AA887131	Nucleotide sequenc
44	50.4	2.6	9700	22	AA890958	Nucleotide sequenc
45	50.4	2.6	9749	22	AA890959	Plasmid pGC1003 nu

## ALIGNMENTS

RESULT	1
AA020567	
ID	AA020567 standard; cDNA; 1912 BP.
XX	
AC	
XX	AA020567;
DT	02-MAY-2001 (first entry)
XX	
DE	Maize RuVB orthologue #2 cDNA.
XX	
KW	Maize; RuVB orthologue; branch migration; heteroduplex extension;
KW	homologous recombination; transformatiion; transgenic plant; ss.
XX	
OS	Zea mays.
XX	
Key	Location/Qualifiers
FT	94..1461
FT	CDS
FT	/tag= a
FT	/product= "Maize RuVB orthologue protein #2"
XX	
FN	W0200105975-A1.
XX	
PD	25-JAN-2001.
XX	
PE	13-JUN-2000; 2000MO-US16271.
XX	
PR	16-JUL-1999; 99US-0144112.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Mahajan PB;
XX	
WP1	WPI; 2001-159537/16.
DR	P-PSDB; AAY72563.
XX	

XX Novel maize RuvB nucleic acid useful for modulating levels of maize  
PT RuvB in plants, as probes or amplification primers in the detection  
PT quantitation or isolation of gene transcripts -  
XX  
PS Claim 1; Page 70-72; 87pp; English.

The present sequence is a *zea mays* RuvB orthologue #2 cDNA. RuvB along with RuvA catalyses the branch migration process, also known as heteroduplex extension, in homologous recombination. RuvB is used for the control of homologous recombination or transformation efficiency in transgenic plants. The RuvB nucleotide may be used as probes or amplification primers for detecting, quantifying or isolating gene transcripts, in detecting deficiencies in the mRNA level during screening for, desired transgenic plants, for detecting gene mutations, for monitoring upregulation of expression or changes in enzyme activity, for detecting any number of allelic variants, orthologues or paralogues of the gene, or for site directed mutagenesis in eukaryotic cells. It may also be used for recombinant expression of its encoded polypeptide, or for use as immunogen in preparing and/or screening of antibodies, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The RuvB proteins may be used in assays to agonise or antagonise the enzyme function, or as immunogens or antigens for screening antibodies.

Sequence 1912 BP; 547 A; 385 C; 515 G; 462 T; 3 other;

Query Match	99.9%	Score 1910.8;	DB 22;	Length 1912;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1912; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	accacgcgtccgcgaattctgttcgcgcgagagacccgagagagagcagctctccaaagaa	60
Db	1	accaccgcgtccgcgaattctgttcgcgcgagagacccgagagagagcagctctccaaagaa	60
QY	61	acagagagcgcataacccgcgcgttcgtgcgcgcgaatgagagatcgaagaggtgcagctcgacc	120
Db	61	acagagagcgcataacccgcgcgttcgtgcgcgcgaatgagagatcgaagaggtgcagctcgacc	120
QY	121	tcgaagaagcagcgcgcatactgcaccaccacacacatacaaggagctcgcctcgaccgaat	180
Db	121	tcgaagaagcagcgcgcatactgcaccaccacacacatacaaggagctcgcctcgaccgaat	180
QY	181	gggacgagcgcattctgcgttcgcgcgcgcgcgttcgttcggccacagccgcgcgcgcgcgcgc	240
Db	181	gggacgagcgcattctgcgttcgcgcgcgcgcgttcgttcggccacagccgcgcgcgcgcgcgc	240
QY	241	gggcctgcgcgtctgcacatgattcgcaccagaagaagaatggccgcgcgcgcgcgcctcttcgc	300
Db	241	gggcctgcgcgtctgcacatgattcgcaccagaagaagaatggccgcgcgcgcgcgcctcttcgc	300
QY	301	ggctccgcgccgcacacgcgcgaagaacgcgcgtcgtacgcctcgcatagccagagcttcgcacgc	360
Db	301	ggctccgcgccgcacacgcgcgaagaacgcgcgtcgtacgcctcgcatagccagagcttcgcacgc	360
QY	361	aaggtccctcttcgtccctatgtagatcacaaagtgtactctctgcgcgcgcgcgcgcgcgcgcgc	420
Db	361	aaggtccctcttcgtccctatgtagatcacaaagtgtactctctgcgcgcgcgcgcgcgcgcgcgc	420
QY	421	gagcgtcgtatgaaataattccgttcgagcgtctaaagtttcgttcgtatataaagagaaacaaagag	480
Db	421	gagcgtcgtatgaaataattccgttcgagcgtctaaagtttcgttcgttcgtatataaagagaaacaaagag	480
QY	481	gtttatgaaagagaggttacttgaactctccacagaagagcgtcgaagatacaactgtgtgga	540
Db	481	gtttatgaaagagaggttacttgaactctccacagaagagcgtcgaagatacaactgtgtgga	540
QY	541	tatgcacaaaagcatatagccatgttaacttcgaacttcgaagactgtttaagggactaaacaa	600
Db	541	tatgcacaaaagcatatagccatgttaacttcgaacttcgaagactgtttaagggactaaacaa	600
QY	601	ctaaacttgcattctccaattatataatgctctgcatacgaagcgaagagctgcgcgttcgtgat	660

Db	601	ctgaagttcagatccctccaaattctatgtatgtctctgcacgaaggaaaggtgycaggtgcygcat	660
Qy	661	gtctatcatctcgaaacnaaatagttgagtcagttgaaagagttggtatgattgtatctctt	720
Db	661	gttatcatctcgaaagaaatagttgagtcagttgaaagagatgtgtatgattgtatctctt	720
Qy	721	gtctacaatatgacgtctcttgaaagtcgaaagagattgttctctatccccaaggtggaagtcac	780
Db	721	gtctacagaaatagcgtctcttgaaagtcgaaagagattgttctctatccccaaggtggaagtcac	780
Qy	781	aagaaaaaagaatgtgtcagagatgtcaaatctctatgacctgtatgcagaatgtccag	840
Db	781	aagaaaaaagaatgtgtcagagatgtcacactctcatgacctgtatgacctgtcagaatgtccag	840
Qy	841	ccacaaggtgcccagaatatattgttcccttatgggcccagatgatbaaaccccgaaagact	900
Db	841	ccacaaggtgcccagaatatattgttcccttatgggcccagatgatbaaaccccgaaagact	900
Qy	901	gaatccacggaanaaacacgcgccagaagaattlaaiaaggtgtgtaaatatagatatatcatgtgaa	960
Db	901	gaatccacggaanaaacacgcgccagaagaattlaaiaaggtgtgtaaatatagatatatcatgtgaa	960
Qy	961	ggaattgcagaagctgtgacctgtgtttgtgtcatgtatgaggtcccaatgtttgatat	1020
Db	961	ggaattgcagaagctgtgacctgtgtttgtgtcatgtatgaggtcccaatgtttgatatc	1020
Qy	1021	gaattttttcttacttaacccgtcatgtaggaagcccatctacccaatgctgatact	1080
Db	1021	gaattttttcttacttaacccgtcatgtaggaagcccatctacccaatgctgatact	1080
Qy	1081	gtctacaaatagggagatactglaatgtlaaagagaagaactgtatagtacaaagctccaaatgtgata	1140
Db	1081	gtctacaaatagggagatactglaatgtlaaagagaagaactgtatagtacaaagctccaaatgtgata	1140
Qy	1141	ccgggtgagctctctcagaatagtcggtgatatcttcggaacgaacatatgcccctactgag	1200
Db	1141	ccgggtgagctctctcagaatagtcggtgatatcttcggaacgaacatatgcccctactgag	1200
Qy	1201	atgatacagatattggtcctatccgagccaagtggaaggaaatctgtatatgtaaagaagaat	1260
Db	1201	atgatacagatattggtcctatccgagccaagtggaaggaaatctgtatatgtaaagaagaat	1260
Qy	1261	cttgcttatttaggcggagatcgcgacacgaacactctttagacatgcatctcaattgtga	1320
Db	1261	cttgcttatttaggcggagatcgcgacacgaacactctttagacatgcatctcaattgtga	1320
Qy	1321	tcacctgcagcggtgctctcaagaactatggaagagaaatctgcgaagctgatactc	1380
Db	1321	tcacctgcagcggtgctctcaagaactatggaagagaaatctgcgaagctgatactc	1380
Qy	1381	gaggaagtcagttggtctctatttgatgtccaaatcctctggctcggctgtctccagagaaga	1440
Db	1381	gaggaagtcagttggtctctatttgatgtccaaatcctctggctcggctgtctccagagaaga	1440
Qy	1441	caagaagaatcatcaccataatttggatctcctgtcgtggaagtcctcgaaagagaaatgla	1500
Db	1441	caagaagaatcatcaccataatttggatctcctgtcgtggaagtcctcgaaagagaaatgla	1500
Qy	1501	gttcgcagctcgaaagttcatctagttgattgtgactgtcttccaaaggttcatatgctcaacgg	1560
Db	1501	gttcgcagctcgaaagttcatctagttgattgtgactgtcttccaaaggttcatatgctcaacgg	1560
Qy	1561	tccttgagcgagacattcttgagggggaagcgcttgaaatttggagagccctgtgtgtla	1620
Db	1561	tccttgagcgagacattcttgagggggaagcgcttgaaatttggagagccctgtgtgtla	1620
Qy	1621	gtctccaragaagaactgtgtctcggacatattgctgttcaagcagatgtctcgtattaga	1680
Db	1621	gtctccaragaagaactgtgtctcggacatattgctgttcaagcagatgtctcgtattaga	1680
Qy	1681	tgtgtcaccggtgcagaagaattgcggtgtgtgtgttttatctgtcatctggttccggaggt	1740
Db	1681	tgtgtcaccggtgcagaagaattgcggtgtgtgtgttttatctgtcatctggttccggaggt	1740





QY 346 caggagctcgccagcaaggctccctctctgctccatgtagatcaagaagtgtactctcg 405  
 |||||  
 Db 265 caggagctcgccagcaaggctccctctctgctccatgtagatcaagaagtgtactctcg 324  
 QY 406 gaggtaagaagaactgaggctgctgatacgaaattccgtagagactaaggcttgatata 465  
 |||||  
 Db 325 gaggtaagaagaactgaggctgctgatacgaaattccgtagagactaaggcttgatata 384  
 QY 466 aaggaaacaagaagaggtttatgaaggagaggttactgaaacttccccgaagaaggctgag 525  
 |||||  
 Db 385 aaggaaacaagaagaggtttatgaaggagaggttactgaaacttccccgaagaaggctgag 444  
 QY 526 agtacaaactggtgatactgcaaaaagcattagacatgtaacatcaagcttaaggctgtt 585  
 |||||  
 Db 445 agtacaaactggtgatactgcaaaaagcattagacatgtaacatcaagcttaaggctgtt 504  
 QY 586 aaggagactcaagcaactggaagttaattcttcaattatgatacgctctgatacaagaagaag 645  
 |||||  
 Db 505 aaggagactcaagcaactggaagttaattcttcaattatgatacgctctgatacaagaagaag 564  
 QY 646 gtgagcaagtgtgatactatatactgcaagaactagtgagagcagtgaaaaaggttgt 705  
 |||||  
 Db 565 gtgagcaagtgtgatactatatactgcaagaactagtgagagcagtgaaaaaggttgt 624  
 QY 706 agatgtgattcttctgtctacagaatacgaacttgaaagctgaagagatggttccctacccc 765  
 |||||  
 Db 625 agatgtgattcttctgtctacagaatacgaacttgaaagctgaagagatggttccctacccc 684  
 QY 766 aagggtgaagctccataagaaaaaaagaattgtgcagagctgcaacactcaatgaccttgat 825  
 |||||  
 Db 685 aagggtgaagctccataagaaaaaaagaattgtgcagagctgcaacactcaatgaccttgat 744  
 QY 826 gcaagcaaatgtctcagcccaagctgagccaagatatgttgcctcattagggccagatagt 885  
 |||||  
 Db 745 gcaagcaaatgtctcagcccaagctgagccaagatatgttgcctcattagggccagatagt 804  
 QY 886 aacccaagaagaactggaatacaccgaaactaagccaagaataataaagtgtgtaaat 945  
 |||||  
 Db 805 aacccaagaagaactggaatacaccgaaactaagccaagaataataaagtgtgtaaat 864  
 QY 946 agatatcgaatggaaggaattgacagagctgtgactgtgtgttgcctcattaggaagctc 1005  
 |||||  
 Db 865 agatatcgaatggaaggaattgacagagctgtgactgtgtgttgcctcattaggaagctc 924  
 QY 1006 cacatgtgtgatalcogaatgttcttctatcttaacgctgacatggaagcccaattca 1065  
 |||||  
 Db 925 cacatgtgtgatalcogaatgttcttctatcttaacgctgacatggaagcccaattca 984  
 QY 1066 ccaatcgtagatactgtctacaaataaggggaatatagttaaggaagaactgatalgaca 1125  
 |||||  
 Db 985 ccaatcgtagatactgtctacaaataaggggaatatagttaaggaagaactgatalgaca 1044  
 QY 1126 agtcacaatgtagatccggtgagatctctagataggtgtgtgtatctcgacaagagaca 1185  
 |||||  
 Db 1045 agtcacaatgtagatccggtgagatctctagataggtgtgtgtatctcgacaagagaca 1104  
 QY 1186 tatggccctactgagatgatalcagatacttgctatccgagacacaagtggagagatgtgat 1245  
 |||||  
 Db 1105 tatggccctactgagatgatalcagatacttgctatccgagacacaagtggagagatgtgat 1164  
 QY 1246 atggatgaagaagaagctctgtctatttaggcgagatccgacagacagacatcttgagacat 1305  
 |||||  
 Db 1165 atggatgaagaagaagctctgtctatttaggcgagatccgacagacagacatcttgagacat 1224  
 QY 1306 gctatcaatgtagatcaactgtagcagcgtgtgtctcaaaagctataggaagagaagaatac 1365  
 |||||  
 Db 1225 gctatcaatgtagatcaactgtagcagcgtgtgtctcaaaagctataggaagagaagaatac 1284  
 QY 1366 tgcgaagctgatatctcgaggaagctcagtggtctatcttgatgccaatccctcgctcg 1425  
 |||||  
 Db 1285 tgcgaagctgatatctcgaggaagctcagtggtctatcttgatgccaatccctcgctcg 1344  
 QY 1426 ctgctcagagagacaagaagaatacatalcaactagatttgatctctctgtcgtggaagt 1485

|||||  
 Db 1345 ctgctccagagagcaacaagaagaatacacccttagatttggctacactgtcgtggaagt 1404  
 QY 1486 ctgcaagaagaatgtagttagtgcagactcgaaagtcaacttagtgcattgtagtcttcacagg 1545  
 |||||  
 Db 1405 ctgcaagaagaatgtagttagtgcagactcgaaagtcaacttagtgcattgtagtcttcacaca 1462  
 QY 1546 ttcatagctacacggtcttgagcgagacacatttcggggggaacggttgaaatttgcaag 1605  
 |||||  
 Db 1463 -----gctcttgagcgagacacatttcggggggaacggttgaaatttgcaag 1509  
 QY 1606 tgcctgctgtgtgtagtctccaragaagacttggttccgcgacatctgctgcagcagac 1665  
 |||||  
 Db 1510 tgcctgctgtgtgtagtctccaragaagacttggttccgcgacatctgctgcagcagac 1569  
 QY 1666 tgtcgcgtgattagattggttcacccggtgcaggaattgcccgtgtgtg--tttlatctgtc 1724  
 |||||  
 Db 1570 tgtcgcgtgattagattggttcacccggtgcaggaattgcccgtgtgtg--tttlatctgtc 1629  
 QY 1725 catcggtgtccggagtc--tgctccacgggtgtgatttgcccggaacctatcttgtaa 1782  
 |||||  
 Db 1630 catcggtgtccggagtc--tgctccacgggtgtgatttgcccggaacctatcttgtaa 1689  
 QY 1783 ccatggaataatggaatggaatcttctacagaatgcaacttgatggcttattatcttaaa 1842  
 |||||  
 Db 1690 ccatggaataatggaatggaatcttctacagaatgcaacttgatggcttattatcttaaa 1749  
 QY 1843 tgtccataaagaacaaagaaatgtttctacaacmtwtaa 1882  
 |||||  
 Db 1750 tgtccataaagaacaaatgtttctacaacatataaga 1789

## RESULT 4

AD023568  
 ID AAD02568 standard; cDNA; 1886 BP.

AC AAD02568;

DT 02-MAY-2001 (first entry)

DE Maize RuVB orthologue #3 cDNA.

KW Maize: RuVB orthologue; branch migration; heteroduplex extension;  
 homologous recombination; transformation; transgenic plant; ss.

OS Zea mays.

Key Location/Qualifiers

FT CDS 82..1449

FT /tag= a

FT /product= "Maize RuVB orthologue protein #3"

PN W0200105975-A1.

PD 25-JAN-2001.

PF 13-JUN-2000; 2000MO-US16271.

PR 16-JUL-1999; 9905-0144112.

PA (PION-) PIONEER HI-BRED INT INC.

PI Mahajan PB;

DR WPI; 2001-159537/16.

DR P-PSDB; AAY72564.

XX

Novel maize RuVB nucleic acid useful for modulating levels of maize  
 RuVB in plants, as probes or amplification primers in the detection,  
 quantitation or isolation of gene transcripts.

PS Disclosure: Page 74-76; 87pp; English.

XX











[illegible]



```

Db      691  gtcccttgcacaaaggagatgctgcacaaagaagaatacctacagatgtgaccttg 750
Qy      814  catgaccttgatgcacaaatgtctcagccacaagctgcccagaatatttctccctatg 873
Db      751  cagcacttgacgtgagccatctgcgcctcaggtgtggcagaatatttctctatgatg 810
Qy      874  ggcacatgataaaccacgaagaagactgaatcaccgaataaatactagccacaataat 933
Db      811  ggcacgttgatgagccacaaaagaacagacagatcacagaataaacttcgaggagatc 870
Qy      934  aaggtgttaataatagatatcagtagaagaattgcagagctgttactcctgtttgttc 993
Db      871  aaggtgtgtaacaaatacattgaccagagcgttgcagagctgttccctgtgagctctt 930
Qy      994  attgatgagctcacatggttggatctcgaatgttttcttacttacttaaccgttcat 1053
Db      931  gttgacgaagttccacatgctgcatatcagatgttacttacttaccctgcacgcagc 990
Qy      1054  agccattatcaccaatcgttgatctgtactgtacaataagagggaatagttaagtaga 1113
Db      991  tctccatcgcccccattgtcatcttggatcaccacagaggaactgtgtcatcagggcc 1050
Qy      1114  ac---tgatatgacaagtcacacatgttatcccggttgatcttctagatagcgttgat 1170
Db      1051  accgaagacatctacttctccacacgcatcccggttgatctgtgaccggtgtgatc 1110
Qy      1171  attcgacagagacatattgcccctactgagatcacagatatgttgctatccgacacaa 1230
Db      1111  atcagagaccatgctgtatcagccacagagatgtagagccagatcatatcaccgacag 1170
Qy      1231  gtgagagagattgatatgtagaagaagctgtcttacttattagcgagatcggacaagag 1290
Db      1171  acggaagacatcaatcagtagaagagccctaaaccacctcggaggagattggaccaa 1230
Qy      1291  acatcttgagacatgcttcatcaatgttatcactcgtccacggtgtgtcacaagaact 1350
Db      1231  accagcctgaggtattctcagtgacgtcgtcgaacccctgcacacctgtgccaagatc 1290
Qy      1351  ggaagagagaatctgcgaagctgtatcgcgaagagtcgaatggtggtctatttgatgcc 1410
Db      1291  ggggaagagacagattgagaagagagacgtgtgagagatcaccagcgtcttctatgag 1350
Qy      1411  aatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1470
Db      1351  aagtcctccgcaagatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1410
Qy      1471  t 1471
Db      1411  t 1411

```

RESULT 8  
AA18199 standard; cDNA to mRNA: 1730 BP.

AC AAX18199;  
DT 10-MAY-1999 (first entry)  
XX TIP49 coding sequence #2.  
DE TIP49 coding sequence #2.  
KW TIP49; TATA binding protein; TBP binding protein;  
KW DNA transcription control; ss.  
XX Rattus sp.  
OS Rattus sp.  
PN M09900419-A1.  
XX 07-JAN-1999.  
PD 25-JUN-1998; 98MO-JF02836.  
XX

```

PR      27-JUN-1997; 97JP-0187398.
XX
XX (SOME ) SUMITOMO ELECTRIC IND CO.
XX Kishimoto T, Makino Y, Niwa S, Tamura T;
XX WPI: 1999-095682/08.
XX P-PSDB: AAW74417.
XX
XX TBP-binding protein with DNA helicase and ATPase activities - gene
XX encoding it, and antibodies recognising it.
XX
XX Claim 10: Page 43-45; 64pp: Japanese.
XX
XX This sequence encodes the TATA-binding protein (TBP) binding protein,
XX designated TIP49, of the invention. TIP49 and its associated nucleic
XX acids and antibodies are useful in investigation of the process of DNA
XX transcription control by TBP in vivo, and in detection of the blocking of
XX DNA transcription.
XX
XX Sequence 1730 BP: 469 A; 401 C; 476 G; 384 T; 0 other;

```

Query Match 32.0%; Score 612.4; DB 20; Length 1730;  
Best Local Similarity 66.0%; Pred. NO. 7.8e-131;  
Matches 902; Conservative 0; Mismatches 461; Indels 3; Gaps 1;

```

Qy      94  atgagatcgaagagagtgatcagctcgaagaagagcgcacatcgccaccacacac 153
Db      60  atgagaattgaggggtggaagagacatacgaagcgcagcgcacatcgccaccacacac 119
Qy      154  atcaagagctcggccctcgcacacgaatgagatggatgttcgttgcgcgggttcgt 213
Db      120  gtgaaagagcggggcgacgagagcggttgcgaagagcggttcgaagcggttcgt 179
Qy      214  ggcacgagcgcgcgcgcgagcgcgcggttgcgttcacatgtttccgacgaagaag 273
Db      180  ggcacgagagagcgcgcgagcgcggttgcgttcacatgtttccgacgaagaaga 239
Qy      274  atggcgcgcgcgcggttgcgttcacatgtttccgacgaagaagaagcggttcg 333
Db      240  atggcttgaagaatgttgcgttcacatgtttccgacgaagaagaagcggttcg 299
Qy      334  ctggcctagccagagagcgcgcgcacgaagcttccttgcgttcctatgtatgataaga 393
Db      300  ctggcctatgtcgaagagcggttgaagaagcttccttgcgttcctatgtatgataaga 359
Qy      394  gtgtactcctcggaggtlcaagaagaatgaggttcatgtagaagaatcgttagatc 453
Db      360  gttactcaactgagatcaagaagaagagaggttcatgtagaagaatcgttagatc 419
Qy      454  gtttcgtataaagaagaagaaggttcatgtagaagaaggttcatgtagaagaatcgt 513
Db      420  ggttcgtataaagaagaagaaggttcatgtagaagaaggttcatgtagaagaatcgt 479
Qy      514  gaagagcgtgagatgatacactgtgtatgatacgaagaagcattagccatgatac 573
Db      480  tgltagacagagatcccatgtagagatgtagaagaatcattagccatgatac 539
Qy      574  ttaagactgttaagaggaactaagcaactgaagttagaattcattatgtagatcgt 633
Db      540  ctcaaaacagccaaaggaacaaacaggttgaacatggaaccagacattttgaaggt 599
Qy      634  atcaaggaaggtggtgaggtgtgtatgatacgaagaagcattagtagaagcag 693
Db      600  cagaagaagcgtgataaagcgtgagatgtagatcattatggaagcacaaggtggt 659
Qy      694  aaaaagattgtgtagatgttcttctgtacagatacgaatcgtgaagcgtgaagat 753
Db      660  aagaagcagggcaggtgtgatacctatgacacagatcgcactggaagcgtgaagat 719
Qy      754  gtctctatcccaaggtggaagtcataagaaaagaagaatgtgtcagagatgcacact 813

```



OY	694	aaagaagttgtatgatactgattcttttgcctacagaatacagatcttgaagcttgaagat	753
Db	682	aagaagcgacgggacgggtttgtatctacttagccacagaattcgaacttgaagtgaaatt	741
OY	754	gtctctatccccaagggggaagtcctatagaanaaagaatttgcagatgtcacaact	813
Db	742	gtcccttcgccaanaaggggagatggtcacaaaagaagaatactacagaatgtgcacttg	801
OY	814	catgaccttgaatgcagcaaaatgcctcagccacaagaatgtgcacaatatcttgcctatg	873
Db	802	catgacttggatgtgtgtctaatgctgcggccccaaggggggaagaatatctctgtccatgatg	861
OY	874	ggccagatgatggaaccacgcaagaagactgaaataccgcgaaaaactacgcagcaagaataat	933
Db	862	ggccagctaatgtgaagccaaaagaacagaaatcacagacacaactctgaggggagaattaat	921
OY	934	aagggtggaataagatatactatcgaatggaaggaattgcagagattgtacctgtgtttgtctc	993
Db	922	aaggtgtgtgaacaaagtatacttcgcacccaagggaatgtgcgtgtgtccgggtgtgcgttt	981
OY	994	attgaatgaagtcacacatgttttgatatacgaattgtttctctatcttaacgctgcattggag	105
Db	982	gttgaatgaagtcacacatgttcgacattgagtgtcttaacctgaactgcagccgcttggag	104
OY	1054	agcccatattatccaatccgtagaactctgtcacaanaatagggggaatagtgaatgaagaga	1113
Db	1042	tctctatccgtctcccatccgtatctatcttgcctaccacccagagcaactgtgtcatcagagcc	1101
OY	1114	act--gatagaagaatgccacatgagatccggttgatctcttgatgatgctgtgtgatt	1170
Db	1102	actgaggaacatcaatactccctcaagcatccctcttgcacccctctgcgaacggagtgtgata	1161
OY	1171	attcgacacagagacatatgcccctactgaatgatatacagatatgttgatctccgagcaaca	1230
Db	1162	atccgagacatgctgtatactaccacaggaatgaacagatcattaaatccgtgtgccag	1221
OY	1231	gttgaggaagatgtatgatgagtggaagaagtctgtctatttgagcgagatccgagacag	1290
Db	1222	acggaagaagaatcaacaacagtgtagaggggaacttgaacacactctgtgggagattgtgcaccaag	1281
OY	1291	acactcttgagacatgtctattcaattgatataccctgcagcggtgtgtctcaagaagctaat	1350
Db	1282	accacacttggtgtactccagtgatgcagctgcgtgcaccccggcacaactgtgtgtcttaaatcaac	1341
OY	1351	ggaaagagagaanaatctgcgaaggtctgactcgcgaagaatgcagtggtgctctatttggatgc	1410
Db	1342	gggaagagacagcatctggaagaagcatgtcgcgaagatcatcgaactttctatgatgcc	1401
OY	1411	aaatcctcgtctcgtctgtctccaggggcaacaagaagaatcatca	1456
Db	1402	aagctccctccgcaaaatccgtgtgtgcacggatgaaatgaatgata	1447

	RESULT	10
AAC59220	ID	
AAC59220	standard;	CDNA; 1744 BP.
XX	AC	
XX	AAC59220;	
XX	29-JAN-2001	(first entry)
DT		
XX	Human secreted protein CDNA sequence #6.	
DE		
XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;	
KW	antiallergic; hepatotrophic; antidiabetic; antiinflammatory; anticancer;	
KM	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;	
KM	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	
KM	neurological disease; infection; human; secreted protein; ss.	
OS	Homo sapiens.	
XX		
XX	WO200055199-A1.	
NN		

XX 21-SEP-2000.  
PD  
XX  
XX PF  
XX 09-MAR-2000; 2000OWO-US06014.  
XX  
XX 12-MAR-1999; 99US-0124095.  
PR 11-JUN-1999; 99US-0138598.  
PR 03-DEC-1999; 99US-0168665.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM, Komatsoulis G;  
PI  
XX WPI; 2000-572359/53.  
XX P-PsDB; AAB27799.  
XX  
PT Isolated nucleic acid molecule encoding a human secreted protein is  
PT used in preventing, treating or ameliorating a medical condition -  
XX  
XX  
XX Claim 1; Page 343; 433pp; English.  
XX  
XX The invention relate to the isolation of genes AAC59215-C59261 encoding  
CC 47 human secreted proteins AAB27794-B27840. The genes can be used to  
CC generate fusion proteins by linking to the gene for the human  
CC immunoglobulin G<sub>1</sub>FC portion for increasing the stability of  
CC the fusion protein as compared to the human protein only. The genes and  
CC proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated  
CC from a range of human tissues disclosed in the specification. The  
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. Breast  
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) Immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, Rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
XX  
XX Sequence 1744 BP; 476 A; 403 C; 480 G; 384 T; 1 other;

Query Match	28.5%;	Score 544.4;	DB 21;	Length 1744;
Best Local Similarity	65.7%;	Pred. No. 2.9e-115;		
Matches 898;	Conservative 1;	Mismatches 457;	Indels 10;	Gaps 7;
QY 94	atgagatctcgagagagtgctgcagtcgcagcttcgaaagaacagcgatcagccacccac	153		
Db 76	atgaagatctgaagagtgtaagagacactacgaagacagcgatcgctccacacagccac	135		
QY 154	atcaagggagctcgctctgcagcccaatgtagtgcgatgtggttcgcgcggttcgtg	213		
Db 136	gtgaagagggcttgaggctcgtgagcagagcggtcttgcgcaagcagcgctccacaggtctg	195		
QY 214	ggccagggcgcgcgcgcgagagcgccggtcggcgctgcgaatgtatccgccaagaag	273		
Db 196	ggccagggagaaacgcgcgagagagcaltgctggtcaltagatgaatlaatcaaaagcaga	255		
QY 274	atggcggcgccgcgcggtctcctcttcgttcggctccgcccgcgaaggaagaacggtacg	333		
Db 256	atggctcggaagagcgtgctcttggtgcgaagacctcctggaactatggaagaacagcttgct	315		
QY 334	ctcgcatagacccgaagactcgcgcagcaagctcccttcgttcctatgtatgatacga	393		
Db 316	ctggtctattgctcagagagctggtgtagtaagttccctcttcgccaatggttggtgagta	375		
QY 394	gtgtactctcggaggttcaagaagaactgaagtgctgtcgaatgaataatccctcaga	453		
Db 376	gttacttcaactggtatccaagaagacagaagtgctgtatgtaggaactcccgag-gccat	434		
QY 454	ggttgcgtataaaggaacaagaaggttlatgaagagagaggttactgaacttccca	513		



QY 650 cagtggtggtatgtatatacagcaagcaaatagtcgagcagtcgaaagagttgtagat 709  
DB 471 aagctggagatgtgtattacattgaagcacaagtgggccgtgaagggagggcaggt 530  
QY 710 gtagtctttgtctacagaatagatcttgaagtcgaagatagttctatccccaag 769  
DB 531 gtagtacctatgcccagaaattcgaccttggaagtcgaagatgtgtccct-tgcctaaag 589  
QY 770 gtagagtcataagaanaaagaattgtcgagatgtacacactacactgactgtatgacag 829  
DB 590 gggatgtgacacaaaagaanaagaatcccaagatgtgacactgacactgactgtatgacag 649  
QY 830 caaatgtcagccacaaagtgccaaagatattgtccctatgtggccagaatgtatgaac 889  
DB 650 ctaatgctggcccccagggggagacagatctccctgatgatgtggccagtaatgtagc 709  
QY 890 caggaagaactgaataccacccgaataactgacccaaagaaataaagtggttaaatgat 949  
DB 710 caaagaagaagcaaatccacagacaaactcgagggagatataaagtggtgaacaagt 769  
QY 950 atatcgaatgaagaattcagagctgtacgtgtgtttgttctatgtagaggtccaca 1009  
DB 770 acatgcaccagagatgtgtgagctgtccgggtgtgctgtgtgtatgtagaggtccaca 829  
QY 1010 tgttgagatcgaatgttttctatctlaacggtgacatgtgagagccatcatcaca 1069  
DB 830 tgcgtgacatgtagtgcttccactacactgacacgcccctggagttctctatcgtccca 889  
QY 1070 tctgtaactgtgtaacaataagggaaatagtaatgaagaagact---gatatgacaa 1126  
DB 890 tctgcaactgtgtaacaacaggaagcaactgtgtacacagaggaagcaactgagacatcat 949  
QY 1127 gtcacacatgtatatacgggtgtagtcttctagatagtcgtgtagatattcggacagagacat 1186  
DB 950 cccctcagcgatccctctgactcttgacccgagtgagatgaataatccggacacatgctgt 1009  
QY 1187 atggcctactgtagatcagaatattggtctacgcagacaaagtgagagatgata 1246  
DB 1010 atactccacaggaatgaataacagatcataaactcgtgccagaggaaggaataaca 1069  
QY 1247 tgaatgagaagaatgtctgtatttagcgagatcgagacagacactctttagacatg 1306  
DB 1070 tcaatgagggagcagtcgaacacactggggagatgtgaccacagacacactgaagttact 1129  
QY 1307 ctattcaattgatatatacctgcagcgtgtgtccaagactaattgaaagaaataatct 1366  
DB 1130 cagtgacgctgctgaccccgccaactgtgtctaaatacaacgggaaggaagacatgtg 1189  
QY 1367 gcaagctgtctcgcaggaagatcagtggtgctctatttgatgccaatccctcgctcggc 1426  
DB 1190 agaaagagcatgtcgagagatcagtaactttctatgtatgccaagctcctccgccaata 1249  
QY 1427 tgcctccagagcacaagaagaatatacactacactgagattgtagtctctgtcgtggaagtc 1486  
DB 1250 tctctgggtcgtgacacagcagggatgaatgaatcattgaagtgtgagtggtttt 1309  
QY 1487 tcgaagaagaatgtagttgc 1506  
DB 1310 cagcagcagtaagagactccc 1329

KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
OS Aspergillus oryzae.  
XX  
PN WO200056762-A2.  
PD 28-SEP-2000.  
XX  
PE 22-MAR-2000; 2000WO-US07781.  
XX  
PR 22-MAR-1999; 99US-0273623.  
XX  
PA (NOVO ) NOVO NORDISK BIOTECH INC.  
PI (NOVO ) NOVO NORDISK AS.  
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
DR MPI; 2000-594572/56.  
XX  
XX  
PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags -  
XX  
PS Claim 88; Page 2561; 3161pp; English.  
XX  
CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring  
CC the global expression of genes from FF cells allows the production  
CC potential of the microorganisms to be improved. New genes may be  
CC discovered, possible functions of unknown open reading frames can be  
CC identified and gene copy number variation and stability can be  
CC monitored. The expression of genes can be used to study how FF cells  
CC adapt to changes in culture conditions, environmental stress, spore  
CC morphogenesis, recombination, metabolic or catabolic pathway  
CC engineering. Using ESTs provides several advantages over genomic or  
CC random cDNA clones including elimination of redundancy as one spot on an  
CC array equals one gene or open reading frame, and organisation of the  
CC microarrays based on function of the gene products to facilitate  
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from  
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus  
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and  
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are  
CC all specifically claimed in the present invention.  
XX  
SQ Sequence 1213 BP; 336 A; 325 C; 275 G; 276 T; 1 other;  
Query Match 17.6%; Score 336.8; DB 21; Length 1213;  
Best Local Similarity 61.8%; Pred. No. 8.7e-68;  
Matches 552; Conservative 0; Mismatches 338; Indels 3; Gaps 1;  
QY 442 cgtagagctataggttgcgtataaagaaacaaagagtttgaagagaggttact 501  
DB 1 cggagagcgatgtgtctcgtcgtcgtgtaaacgaaagagtgatcggaaggagtcacc 60  
QY 502 gaacttcccagaagagcgtagatcactgtgtgatatgcaaaaagcattagccat 561  
DB 61 gagcttaccctggaggaactggaatctcattggagatgtgacgcacacatcagttac 120  
QY 562 gtaatcactcaagcttaaaagcttaaaaggaactgaagctgaagttagattctcaatc 621  
DB 121 ttgattatgttagtaagtcgtcgaaggaacacagaagctgcgtctgtatccagcaat 180  
QY 622 tatatgctctgatacagaagaagtggtgagtgatgttatatatacactgaagaacat 681









```

Db 436 cagattgacgaccgaacacgagcggtcccaaggctgggcaactgcacccctaagacc 495
Oy 568 atcagcttaagactgtttaaggactaagcaactgaagttatgtcttcaattatgat 627
Db 496 acagagatgagacatcacg-----acctggcaccacaagatgatgag 540
Oy 628 gctctgatacgaagaaagtgtagcagtggtgtagtattatatacgaagcaaatagtgga 687
Db 541 tccctgaccacgaagcagcgggagcgtgacgtatcaacacgacgaagcgagcgagc 600
Oy 668 gcaatgaaagagtgtagatgtatcttctgctacagaatacgaatctgaaag----- 742
Db 601 aagatctccacagctggcgctctctcaacagcgcccgagctacgacgtatggctcc 660
Oy 743 -ctgaagagtggttctctatcccaaggtgaagttccataagaagaaagaaatgttcag 801
Db 661 cagaccagagtcgtgacgtgcccagatgggagctccagaacgcaagggatggtagac 720
Oy 802 gatgcacacttcacgtacgtacgacgaatgtctcagccacaaggtggcgaagatatt 861
Db 721 accgtgtccctgcagacgagatcgacgtacatcaactcgcacccagggcttctggcgtc 780
Oy 862 ttgtcccttatggccagatgataaacacgaaagactgaatcaccgaaaaactagc 921
Db 781 ttctc-----aggtgacacagggagatcaagttcagaagttcgt 819
Oy 922 caagaataataagtggttaatatagatatcgatgaagaaatggagaagctgtact 981
Db 820 gacgagatcaatgcgaaggtgtgagtgagcgagagggcgaagggcgaatcattccct 879
Oy 982 ggtgtgttgcattgagatgagacacatgtgtgatacgatgatttcttactacta 1041
Db 880 ggaagcgtctcatcagagaggtccacactgtgtgacacacgagagcttctctccaaac 939
Oy 1042 cgtgattgagagagccattatcaccaatcgttactgtctactaataagggaaatgt 1101
Db 940 cgggcctcgagatgagatgagcgttctctgatacgtgcacacaacggtgacacg 999
Oy 1102 aatgtaagaggaactgatatgacaagtcacatgtatcagcggtgagatcttcatag 1161
Db 1000 cgaacccgggacacagctaccagagccctcagcgacaccccaatagacgtctgagc 1059
Oy 1162 ctgtgattatcggacgagacataltgacctactgagatgacagataltgtgtatc 1221
Db 1060 ctgtatcgtctccacacccctcacagcgagaaagacagacgagatctctcgatc 1119
Oy 1222 cgagcacaagtgagagagatgatatgataagaagaatcttcttatttggcgagatc 1281
Db 1120 cggtgcgaggaagaagatgtgagatgagtgagcgcctacacaggtgtgacccgac 1179
Oy 1282 gacagcagacacatttgaagacatgtctatcaattgatatcacctgcgacggtgtca 1341
Db 1180 gggctggagagagctacgtcgtacgcatccagctcacaacagcttgcaggtgtg 1239
Oy 1342 aagactaatgaaagagaaatctgcagagctgactcgaaggaagtcagtggtctat 1401
Db 1240 cggaaacgcaaggtgacaggaagtgagtgagtgagcaacagcggttactactcttc 1299
Oy 1402 ttgagtgcaaatcctgctgcgtgctgtccagagacgaacgaagaagatacacta 1460
Db 1300 ctggaagagctcccgctccacgacagtaacatgaagagatccagagccttctctaa 1358

```

## RESULT 15

AAAF74190

ID AA74190 standard; DNA; 1280 BP.

AC AA74190;

XX 02-MAY-2001 (first entry)

DE DNA encoding environmental stress tolerant protein SEQ ID 7.

```

KW Environmental stress resistance; salt; heat; desert; transgenic plant;
KM ds.
OS Bruguiiera sezanquila.
PN WO200106006-A1.
XX 25-JAN-2001.
XX 19-JUL-2000; 2000WO-JP04862.
XX 19-JUL-1999; 99JP-0235910.
XX 24-MAR-2000; 2000JP-0085377.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA Yamada A, Ozeki Y, Saito T;
PI
XX
XX
XX MPI; 2001-147355/15.
DR P-PDB; AAB80611.
PT Screening method to obtain DNA encoding environmental stress resistance
PT factor, useful for producing transgenic plants resistant to
PT environmental stress -
XX Claim 23; Page 84-86; 167pp; Japanese.
XX
XX Polynucleotide sequences AA74187 - AA74218 encode proteins
CC AAB80608 - AAB80639, which impart environmental stress resistance. The
CC invention relates to a method for identifying DNA encoding proteins
CC imparting environmental stress resistance. The method comprises inserting
CC cDNA from a library originating in a salt-resistant organism into a host
CC cell, culturing the transformants under conditions in which the
CC untransformed host does not grow well, and selecting for viable clones.
CC The method is useful for obtaining DNA encoding environmental stress
CC resistance factors. The DNA encoding proteins conferring environmental
CC stress resistance, can be used in the production of plants resistant to
CC environmental stress, which can be cultivated in unfavourable
CC environments such as deserts, salt damaged ground, cold regions and the
CC oceans. They can be used for increasing the area of land covered by green
CC plants, and desert greening and afforestation, in order to counter the
CC effects of the increase in atmospheric carbon dioxide concentration. PCR
CC primers AA74219 and AA74220 are used in an example illustrating the
CC method of the invention.
XX
XX Sequence 1280 BP; 388 A; 231 C; 344 G; 317 T; 0 other;

```

Query Match 8.7%; Score 166; DB 22; Length 1280;

Best Local Similarity 51.9%; Pred. No. 1,1e-28; Matches 444; Conservative 0; Mismatches 385; Indels 27; Gaps 2;

```

Oy 605 agttagatttcaattatattgatgctctcatcagaagaaagtgagtgatgtgta 664
Db 110 atttgggggggaaatagatagagcatgtgggaaagaaagtgagtgatgtgta 169
Oy 665 tataatcgaaagcaaatagtgagcgagtgaaagagtgtagatgttcttgccta 724
Db 170 ttgcaattgcaagagcgctccgcaaaatatacaagcttggcggttcatcttcgaggtca 229
Oy 725 cagaatagat-----cttgaagctgaaagagatgttcttctcccaaggtgaagtc 778
Db 230 gggattacgatgcatgtgggacacaggtgaagttgttcaagtgccctgttgggagctgc 289
Oy 779 atagaanaaagaatgtgcaagatgtcacacttgaacattgacattgagcaaatgctc 838
Db 290 agaagaggaagagtgctgcatgtgtctcactcagcagagatgtagtatacaga 349
Oy 839 agccacaagtgagcagaatatttgccttatgagccagatgatacgaacgaaga 898
Db 350 gaacacaggggttcttcttccacccgggagat-----actg 388
Oy 899 ctgaataccagcaaacctacgcgaagaatataagtggttaatatagatatatgatg 958

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2001, 09:53:20 ; Search time 2536.09 Seconds  
(without alignments)  
11661.391 Million cell updates/sec

Title: US-09-589-510-3

Perfect score: 1912  
Sequence: 1 acccaagcgtccgcaattt.....aaaaaaaaaaaaaaaaaaaaa 1912

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 268314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_p11:\*  
13: gb\_p12:\*  
14: gb\_p13:\*  
15: gb\_p14:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_p1:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_hc91:\*  
61: gb\_hc92:\*  
62: gb\_hc93:\*  
63: gb\_hc94:\*  
64: gb\_hc95:\*  
65: gb\_hc96:\*  
66: gb\_hc97:\*  
67: gb\_hc98:\*  
68: gb\_hc99:\*  
69: gb\_hc10:\*  
70: gb\_hc11:\*  
71: gb\_hc12:\*  
72: gb\_hc13:\*  
73: gb\_hc14:\*  
74: gb\_hc15:\*  
75: gb\_hc16:\*  
76: gb\_hc17:\*  
77: gb\_hc18:\*  
78: gb\_hc19:\*  
79: gb\_hc20:\*  
80: gb\_hc21:\*  
81: gb\_hc22:\*  
82: gb\_hc23:\*  
83: gb\_hc24:\*  
84: gb\_hc25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_ro1:\*  
95: gb\_ro2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1910.8	99.9	1912	10	AX077226	AX077226 Sequence
2	1686.6	88.2	1845	10	AX077224	AX077224 Sequence
3	1630	85.3	1869	10	AX077232	AX077232 Sequence
4	1362.2	71.2	1886	10	AX077228	AX077228 Sequence
5	1351.8	70.7	1898	10	AX077230	AX077230 Sequence
6	827.8	43.3	1755	14	CAR276264	AJ276264 Cloner at
7	631.4	33.0	1638	94	BC004718	BC004718 Mus mus
8	631.4	33.0	1709	94	AF100694	AF100694 Mus mus



```

Db 901 GAATACACGAAAACTACGCAAGAAATTAATAGTGTAAATAGATATATCATGATAA 960
Qy 961 ggaattcagagctgttacctgtgtttgttcattatgaggtccacatttgatctc 1020
Db 961 GGAATTGCAAGAGCTGTGACCTGCTGTTGTTGATTCATGATGAGTCCACATGTTGGATATC 1020
Qy 1021 gaaatgtttctatcttaccgttgatcttgagagccattatccaatgttactct 1080
Db 1021 GAATGTTTCTTATCTTAACCGTGCATTCGAGAGCCATTATCAACCAATGATGATCTT 1080
Qy 1081 gctacaaataagggagatatgttaagttaagaggaactgatalgacaagtccacatgtata 1140
Db 1081 GCTACAAATAGGGGAATATGTATTAAGAGAACTGATATGACAAAGTCCACATGATATA 1140
Qy 1141 ccggtgacctcttagatagagctgtgtatcttcggaagaagaacatatgcccactagag 1200
Db 1141 CCGGTGACCTCTTCTAGATAGGCTGTGATATATGCGACAGACATATAGCCCTACAGAG 1200
Qy 1201 atgatacagataatggtctatccagacaaagtggagagatgatalgtatgaagaagt 1260
Db 1201 ATGATACAAATATGCTATGCTATCCGACACAAAGTGGAGAGATTGATATGATGAAGAAAGT 1260
Qy 1261 ctgtcttttttagggagagatcggagacagacacatttgagacatgtctattcaattgata 1320
Db 1261 CTGTCTTATTTAGGCGAATCGGACAGACACATCTTTGAGACATGCTATTCATTCATGATA 1320
Qy 1321 tcaactgcagcgttgctcaagaactaatgagaagaagaataatctcaagctgtatctc 1380
Db 1321 TCACCTGCGCAGGCTGTCTCAAAAGCTATATGAGAGAGAAATCTGCAAGGCTGATCTTC 1380
Qy 1381 gaggaaagtcagtgagctctatttgatgacaaalcctcctcggctcggctccagagca 1440
Db 1381 GAGGAAGTCAGTGGGCTCTATTGATGGCCAAATCCTCGCTGGCTGCTCCAGAGAGCAA 1440
Qy 1441 caagaagaatcatcactatgatttgatctcctgtcgtggaagctcgaagaagtata 1500
Db 1441 CAAGAAGATCATCTACCTATGATTTGATCTCCGTCGTGGAAGTCCGAGAGGAATGTA 1500
Qy 1501 gttgcagctcgaagaatcatctatgtcatgtatcgtcttcacaggttcataagctactag 1560
Db 1501 GTTGCCAGCTCGAAGATCATCTATGATGATGATGATGATGATGATGATGATGATGATG 1560
Qy 1561 tcttgagagcagacattcggggggaaggcgttgaaatttgacagtcgctgtctgtgata 1620
Db 1561 TCTTGAGACCGAGCAATTTGGGGGGGAACGGCTTGAAATTTGCAAGTGGCTGCTGGTGA 1620
Qy 1621 gttccacaagaagactgtgttcggacatctgtgttcaagcactgtctcgtatgata 1680
Db 1621 GTTCACARAGAAAGACTTGGTTCGGGCTATGCTGCTTCACGCACTGCTGCTGATGATA 1680
Qy 1681 ttggtcacccggtgcaagaattcgcgtgtgtgttctatcttgatcgtatcggtgcgagt 1740
Db 1681 TTGCTCACCCGCTGCAAGAAATTTGCCGTGTGTTTATCTTGCTGCTGCTGCTGCTGCTG 1740
Qy 1741 ctgcttcacaggtgtgtatgtgcccgaacccctatcttgaacatgataatgataag 1800
Db 1741 CTGCTTCACAGGGGTGTGATGTCGCGCAACCTATCTTTTAAACATGAGTAATGATAGG 1800
Qy 1801 attctlacagaatgcaactgtgacttatattctaaatgccccataagactaacg 1860
Db 1801 ATTCTTACAAATGCAACTTGGCATGCTTATTTATTTCTAAATGTCATTAAGGCAATACG 1860
Qy 1861 aaatgtttctacaacmwtataaaaaaataaaaaaataaaaaaataaaaaa 1912
Db 1861 AAATGTTTCTTACAAACMTWTAAAAAATAAAAAAATAAAAAAATAAAAAA 1912

```

```

RESULT 2
LOCUS AX077224 1845 bp DNA
DEFINITION Sequence 1 from Patent WO0105975.
ACCESSION AX077224

```

```

VERSION AX077224.1 GI:13121818
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
/organism="Zea mays"
/db_xref="taxon:4577"
85..1452
/feature="unlabeled protein product"
/codon_start=1
/protein_id="CAC32388.1"
/db_xref="GI:13121819"
/translation="MREIYQSTSKKQRLATHTHIKGLIDANGMALAAGFVGSOA
AREAGLAVDMIRKMKAGRAVLVGPATGKTALALGIAOELGSKVPCPVGSFV
SSEYKTEVLMENPRRATIGLRKEVVEGTELSPEASTTGTAKSISHVITS
LKYTKGKOLKIDSSIIDALIKERVAVDVYTHIANSVAVRKVRGCDSPATEYDLAE
EYVTPRGVEVHKKEIVQDVTLDHLDNAANAQPGGDILLNGOMMKPKETETKLR
QELNKVNRITIDGIEALVPGVLEIDVHMLDIEFCSYINRALESPLSVILATNRG
ICNVKGTMSPHGIPVDLIDRLVIRFETGTEMTIOLIAIQAVIDEDIMDESLAY
LGEIGQOTSLRHAIIQLISPAVSYSKTNREKICKADLEVSGLIYDAMSARLLQEQ
ERYIT"
BASE COUNT 530 a 378 c 490 g 445 t 2 others
ORIGIN

```

```

Query Match 88.2%; Score 1686.6; DB 10; Length 1845;
Best Local Similarity 96.5%; Pred. No. 3; 5e-304;
Matches 1758; Conservative 3; Mismatches 42; Indels 18; Gaps 3;

```

```

Qy 46 ggcagctccacagaagaagcgaataacggcggtgttgccgagatgaatcga 105
Db 37 GGCAGCTCCACAGAAACAGAGAGGCAVMAACCGCGCTTGGCGCGATGAGATCGAG 96
Qy 106 gaagtgcagctcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 165
Db 97 GAGGTGCACTGCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156
Qy 166 ggcctcagcccaatgagatgcatgctgttgccgaggtgttgccgagagcgc 225
Db 157 GGCCTCAGCCCAATGAGATGCGATGCGTGGCGCGGGGTGCGCCAGTCCGCG 216
Qy 226 ggcgcgagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 285
Db 217 GCGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
Qy 286 ggcgtgtccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 345
Db 277 GCGGTGCTCTTGTGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 336
Qy 346 caggagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 405
Db 337 CAGGAGCTCGGAGCAAGGTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
Qy 406 gaggtaaaagaagaagtgatgctgataagaaatccgttagagctataggttgcata 465
Db 397 GAGGTCAAGAAACCTAGGCTGATGAGAAATTTCCGATGAGCTATGAGTGTGATATA 456
Qy 466 aaggaagaagaagaagttatgaagagagagtgtaacttcccgagaagaagcgtag 525
Db 457 AAGGAAGAAAGAGAGTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
Qy 526 agtaaacactggttgatataagcaaaagcatagcgaatgaatcaatgaactgaact 585
Db 517 AGTCAACCTGCTGATATGCAAAAGCATTTAGCCATGATATCATCATGCTTAAAGACTGT 576

```





LOCUS	AX077228	1886 bp	DNA	PAT	22-FEB-2001
DEFINITION	Sequence	5	from Patent WO0105975.		
ACCESSION	AX077228				
VERSION	AX077228.1	GI:13121822			
KEYWORDS	.				
SOURCE	Zea mays.				
ORGANISM	Zea mays.				
REFERENCE	Okazaki: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta				
AUTHORS	Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;				
TITLE	Panicloideae; Andropogoneae; Zea.				
JOURNAL	1 (bases 1 to 1886)				
FEATURES	Mahajan, P. B.				
source	Maize orthologues of bacterial ruvB,cdnas and uses thereof				
	Patent: WO 0105975-A 5 25-JAN-2001;				
	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
	Location/Qualifiers				
	1..1886				
	/organism="Zea mays"				
	/db_xref="taxon:4577"				
	82..1449				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC32390.1"				
	/db_xref="GI:13121823"				
	/translation="MRIEEVOSTSKORIATHTHIGLIDANGSMPLAIFVGQAA				
	AREAGLAVDMIROKMKWAGALLAGPAPATGKALALAGTAEISKVPKMGSEVY				
	SSEVKTPEVMEENRRAIGLRIRKENEVYDGEVIELSPREASTTGTAKSISHVITG				
	LKYAKGKOLKLDPSITDALIKERVAGDYIITEANSQVNRKGCSDSFATIEYDEAE				
	EYVLPKEVHKKEIYQDVTLHDLDNANAQPGGDILSLMGQMKRKEITEKLEAK				
	QELKNVNR1IDESIAGLVGVLFIDEVHMLDECHSYNRALESPLIVILATNRNG				
	ICNNRGKNSPHGIPVDLDRLVITIRFETGTEMIQTLARQVEIDIDESLAV				
	LGEGIOQTSIRHAIQILSPASVAVAKTNGREKMKADLEVSGLYLDAKSARLLQEQ				
	ERYIT"				



QY	738	tggaagctgaagagtagtgcctccatcccccaaggtgaagctcacaagaanaaagt	797
Db	813	TCGAAGCTGAAGAGTATGTTCTCATCCCAAGGGTGAAGTCCATTAAGAAAAAGCAATAGT	872
QY	798	gcagaga tgcacaccltcacacttcagacttcgtagcagcaaatgcacgacaaagtggccaga	857
Db	873	GCAGAGATGTCACACTCCATGACACCTTGATGCGACCAATGCCAGCCACAAAGTGGCCAA	932
QY	858	tatttgcctcctatggtggccagatgatagaacccaagaagactgaaatcaaccgaaaact	917
Db	933	TATTTGTCTCCCTTAATGAGCCAGATGATGAACCCACGGACAGATGAATATCCGAAAGT	992
QY	918	acgcagaagaatatlaaagaagtggtaaagaatatatcgatggaagaaattgcagacttgt	977
Db	993	ACGCCAGAAATCAATAAAGGTGGTAAACAGATATATGACGAAGGAATGCGAAGCTTGT	1052
QY	978	accctggtgtttgttcacatctgaatgcacatgtttgacatcgaaatgtgttctctatct	1037
Db	1053	ACCTGCTGTTTGTTCATTGATGAGGTCCACATGTTGATATGGAATTCCTTTCTTATCT	1112
QY	1038	taaccgtgcatcttggaagagcccatatataccaactcgtgatacttgtcacaatataggynaat	1097
Db	1113	TAAAGCTCATTTGGAGAGCCCATTAATCACCAATTTGATACCTGCTACCAATAAGAGAAAT	1172
QY	1098	atgtaatgtaagaggaactgtatagaacatgccaatgacatgataacgggtgactctctga	1157
Db	1173	ATGTAATGTGAGAGAACCCGATATGACGATCCACATGATATACGATGACGATCTTCTA	1232
QY	1158	taggctgtgttattatcggacagaacataatggtccctactgagatgatacagatatggc	1217
Db	1233	TAGGTTGATGATTTATTTGGACAGCAACATATGAGCCCTACTGAGATGATACAGATCTGGC	1292
QY	1218	tatccgagacaaatggaagagatctgatagataagaagaatcttgacttattaggcga	1277
Db	1293	TATCCGACACAAATGGAGAGATGATGATATGATGAAAGAACTCTTGCTTATTTAGGGGA	1352
QY	1278	gatcggagaagacacatccttggagacatgctatcaatgatatacaactgcgaacgttgt	1337
Db	1353	GATCGAGACGACGACATCTTTGAGACATGCTATTTCAGTTGCTATACCTGCGCAGGTGT	1412
QY	1338	ctcaaaagcctaaatggaagagaagaatctgcgaagctgtgctcggagaaagtcagtggct	1387
Db	1413	CGCAAAAGCCAAACGGAGAGAAAAGATGTGCAAGGTGAGCTCGAGGAATGACCGGGCT	1472
QY	1388	ctatttgatgcacaatcctcctggctcgtgcgtgclccagagagcaacaagaagatacalcaac	1457
Db	1473	CTATTGTGATGCCAAATCCCTCGGCTCGCTCGCTCCAGAGGACAAAGAAATACATCAC	1532
QY	1458	ctgaatttgatcctccgtgtcgtggaagctctcgaagaagatgagtgtccagctcgaaagt	1517
Db	1533	CTAAGCTTGCATCTCCTCGCTGCTGGAAGGAAGAGCCGAGAAGAAATGCATTTCGAA--	1590
QY	1518	catctagtgatcatgatactgcttccacaggtgtcatctacgtgctcttggacgagacat	1577
Db	1591	----AATGCGTGCATCTCTTCAGATATGCTAGTCTTCAAGAGCTTGGAGAGAGACGAT	1645
QY	1578	tttggggggg-----aaagcgttgaattttgcagtgccctg-cttgtgtagtctccara	1629
Db	1646	TTTCGCGGGAGAAAGAAATCGTTTGAAATTTGTATATCTCTTTTGTTAAGCTCCAGC	1705
QY	1630	gaagacttggtlccggacatatgtcgtatcgaacgactgtcgtcgattaga---tttgt	1685
Db	1706	GAAGACTTGTATCCGSCATATTTGCTTTGACGACGCTTAAACCCATTAAGATTGGCC	1765
QY	1686	caccggtgcaagaaattgcgcgtgtgtgtttta---tctgtcactcgggttccggagct	1742
Db	1766	CACCGGTGCAAGGAATTCGGGTGTGTTTACCCCTCCGCTTAATCATATGCCAGATCT	1825
QY	1743	gacctcagaggtgtatgtgcccagacctaactcttcttgaacatgagata	1791
Db	1826	GTGTGCTCAAG--GGTGTGATTTAGACCTTACCTTGTAAACATGACACA	1872

RESULT 6  
 CAR276264 1755 bp mRNA PLN 06-MAR-2000  
 LOCUS Cicer arietinum mRNA for putative Ruv DNA-helicase.  
 DEFINITION AJ276264.1 GI:7208770  
 VERSION AJ276264.1  
 KEYWORDS putative Ruv DNA-helicase.  
 SOURCE Chickpea.  
 ORGANISM Cicer arietinum  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Cicer.  
 REFERENCE 1 (bases 1 to 1755)  
 Duplico, B., Esteban, R. and Labrador, E.  
 A RuvB DNA-helicase like protein is expressed in chickpea epicotyls  
 JOURNAL Unpublished  
 TITLE 2 (bases 1 to 1755)  
 REFERENCE Labrador, E.  
 Direct Submission  
 Submitted (29-FEB-2000) Labrador E., Dpto. Fisiologia Vegetal,  
 Univ. Salamanca, Campus Miguel de Unamuno. Pza. Doctores de la  
 Reina s/n, E-37007, SPAIN  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..1755  
 /organism="Cicer arietinum"  
 /cultivar="Castellana"  
 /db\_xref="taxon:3827"  
 /country="Spain"  
 /dev\_stage="seedling"  
 /tissue\_type="reticulated epicotyls"  
 /clone\_id="CAN-5"  
 /note="age 5 days"  
 207..1583  
 /product="putative Ruv DNA-helicase"  
 /protein\_id="CA876908.1"  
 /db\_xref="GI:7208771"  
 /translation="MEKMKIEVOSTTKKQVATHTHIKGLIEVSGKALPFGSVG  
 QAEARFACGLVMDI ROKKMAKRALAGPGGKRALAGTCOELGTRKPRPMVGS  
 EYSTEVEKTEVLENERRAIGLRIKENKVEYEGEVETLSPEETESITGYSISIV  
 IIGKTATGKQLADPTIYDALIKEKVAADVIYEANGAVKRVGRSDRPFTEEDL  
 EAEYVLPKGEVKKKEIVODVTLHDLDAANAPOGGODILSLMGMMRPFTEEDL  
 KLROEINKVNRVYDEGVAEVPELVLFIDEVHMDMCFEYLNALESLSPIYIFAT  
 NRGICTVAGDMTSPHGPYDLDRLVITIRQYGAPEIIOLAIAROVEVLEDEES  
 LAFGEIGQRTSLRHAVQLSPASVAKINGRDNICADLEICSLYLDKSSAKLLQ  
 EQERKYS"  
 BASE COUNT 525 a 332 c 410 g 488 t  
 ORIGIN

Query Match 43.3%; Score 827.8; DB 14; Length 1755;  
 Best Local Similarity 75.1%; Pred. No. 2.3e-144;  
 Matches 1033; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

Db 448 CCTTAGCAGCTTGGGATATGTCAGAGCTTGGACCAAGGTTCCATTGTCCTCAATGCTAG 507  
 Qy 386 gatcgaagtgtactctccggaagtgcaagaactgaggtgtcgtatggaatctccga 445  
 Db 508 GTTCAGAGTATATTCACACAGAAAGTAAGAACCGAGGCTTATATGGAATTTTGAC 567  
 Qy 446 gagctatagttgtcgtataagaanaaagaagttatgaagaaggttactgaac 505  
 Db 568 GGGCTATTGGTCTACGTTTAAAGGAAATTAAGAGGTATATGAAGGCGAGTGACTGAC 627  
 Qy 506 ttccccaagaagagctgaaggtacactgtgtatgacaaagaagctagccatga 565  
 Db 628 TCTCTCCGGAAGAACTGAGACTTTAACAGAGGTTATGGTAAAGTATTAAGCCACTGA 687  
 Qy 566 tcatcagcttaagaagcttlaaaggaactaaagcaactgaatgatcttcaattatg 625  
 Db 688 TAATGGGTTGAAGACAGTTTAAAGAACCAACCAAGCAACTTAAGTTGACCCACGATATG 747  
 Qy 626 atgtctgtatcaagaagaagtgtgcaagtgtatgtatataatcgaagaatagtg 685  
 Db 748 ATGCCCTTGAATTAAGGAAAGTAGCTGTGGGAGTATATATTTAGGCAATATACG 807  
 Qy 686 gaagcagtgaaagagttgtatgtatgtcttctgtacagaagaatgacatctgaagctg 745  
 Db 808 GGCGCTGGAAGAGGTTGGCCAGAGTATGCTTTGCTACTGACTTTGACCTTGAACGAC 867  
 Qy 746 aagaagatcttccatccccaagtgaaagtgcaatcataagaanaaagaatgtgcaagtg 805  
 Db 868 AGAGATATGTTCCACTTCTTAAGGAGAGGTTACAAAAAAGATTTGTCAGATG 927  
 Qy 806 tcaacttcatgacacttgatgcaagaatgtcgaagcacaagtggtccaaagtatttgt 865  
 Db 928 TAACCTACATGATCTGTGATGCTGCCACGACGACTTAAGGGGGCAGATATTTGT 987  
 Qy 866 cccttatggccagatgaatgaacccaagaagctaaatccaccgaagaactgccaag 925  
 Db 988 CTGTTATGGGCCAGATGATGAGCCCTAGGAAACAAATTCATGCAAGTTGAGCAAG 1047  
 Qy 926 aaatlaaagtggtlaaataagatatatcgaatgaagaattgcaagactgtgacctgtg 985  
 Db 1048 AATATTAAAGTTGCAACCGAATATATGAAAGTGAGAGGCTTGACGCTGCTCGGGG 1107  
 Qy 986 ttgtgtcatgtatgaggttccacatgttgatatacgaaagtttttcttcttaaccgtg 1045  
 Db 1108 TTTCTATTATTAATGAGGTGCATATGCTTGTATATGAAATGTTTTTCTTAAATCGTG 1167  
 Qy 1046 catgtgagagccaatatacccaactgtgatactgtctacaaataggggaataatgtaatg 1105  
 Db 1168 CTTTAGAGAGCTCCCTATCTCCCAATAGTATCTTTGGCACAAATTAAGGAATATGACCG 1227  
 Qy 1106 taagaagaactgatatagaagaatgccacatgttataccggtgtgacttctagatagcttg 1165  
 Db 1228 TAAAGGAGACGATATGACCAAGTCTCACGCACTACCTGTTGACTTATGATGATGATG 1287  
 Qy 1166 tgatatattggaagagacatatgtgcccctactgagatgataacagatatgtgctacogag 1225  
 Db 1288 TGATCATTTGCAACCAACTATGCTCGGCTGAATAATATCAAAATTTTATAGCTATCGAG 1347  
 Qy 1226 cacaagtgaagaagatgtatatagaatgaagaagctgtctatctttagcgaagatcgagac 1285  
 Db 1348 CTCGAATGAGAGAACTGTTGTGACGAGGAAAGTTAGCTTTCCCTTGAGGGAATTTGAC 1407  
 Qy 1286 agcagaactcttgaagacatgtctatgaatgtatataaccctgacaggtgtgtccaaaga 1345  
 Db 1408 AACGACCTCCTAAGGATGCTGTTCAGCTTTTATACCTGCTAGTGTTGTGTGCCAAGA 1467  
 Qy 1346 ctaatgaagaagaanaaactgcgaagagctgtatctcgagaagtaagtggtgctattatgg 1405  
 Db 1468 TAAATGAGACGAGACAACTATGCAAGGCGCATCTTGAGGGAATTTGTTCCCTTATATTGG 1527  
 Qy 1406 atgccaaactcctggtcgtgctgtccagaggaacaagaagaaglaacataccta 1460







	Best Local Similarity	66.2%	Prod. No.3.6e-106;	
Matches	914;	Conservative	0;	Mismatches 464; Indels 3; Gaps
QY	94	atgaagatcgaagagagtgacgtgcacgtcgaagaagaagcgcatacgcacacacacac	15	
Db	18	ATGAAGATTGAGAGGTGAAGACACACACCAACGACCAAGCTGGCTCCACACACAC	77	
QY	154	atcaagagacccgagccctcagcgcgaatgagatgagattgcttgccgagcgaggtcgtg	21	
Db	78	GTGAAGGGGGCTGGGGCTGGAGTGAAGAGCGCGCTGGCCAAAGCAGCGCGCTTCGGGGCTCGTG	13	
QY	214	ggccaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag	27	
Db	138	GGCCAGGAGAAACCGGACGAGAGGCATTGTGTGTATAGTCAATTATTCAAAAGCAAGAAA	19	
QY	274	atggccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag	33	
Db	198	ATGGCTTGAAGAGCTGTCTTGTGGCAGGGCGCTCTGGAACCTGGCAAGCAGCGCTTGGCGC	25	
QY	334	ctcgagcatagcccaagagacccgagcaagagctccctctctgctcctatgtagatcaagaa	39	
Db	258	CTGGCTATTGCTCAAGAGACTGGCGAGTAAAGTCCCTTTCGGCCGATGGTGGTGTGTA	31	
QY	394	gtgtactcctcgaaggtcaagaanaatgagtgagtgatgtgaaatctccgtatgagctata	45	
Db	318	GTATTACTCACTGATGATCAAGAAAGACAGAGGTGTATATGAGAACTTCGAAAGGCTATT	37	
QY	454	ggtttcggtataaagaagaacaaagaagtttatbgaagagaggttactgnaattcccca	51	
Db	378	GGGCTGGCGATTAAAGAGACTAAAGGGGTTTTGAAGGGGAGGTGACAGAGCTCACTCC	43	
QY	514	gaagagcgctgaaggtacaactggttgatatactgaagaagattagcagatataatacgc	57	
Db	438	TGTGACACAGACAGACCCCATGGGTGGGTATGCGCAAAACTATAGCCACGTGATCAAGG	49	
QY	574	ttaagaactgtttaaagygactaagcaactggaagtttagattcttcaattatgatatgctcg	63	
Db	498	CTCAAGACTGCCAAAGGAGCAACCAAGCTGAAGCTGACCCGAGTATTTTGAAGATTG	55	
QY	634	atcaagagaaagtgagtcagtggtggtatgtatatacaltcgaaagcaatagtggagcagtg	69	
Db	558	CAGAAAGAACGATGAGGCTGGAAGATGTGATTTCATTGAAGCAAAACAGTGGAGCTGTG	61	
QY	694	aaaagagttgtgtagatgtagattcttcttgtaacagaaatcgctcttgaagcgtgaagagat	75	
Db	618	AAGAGGCAAGGAGGTGTGACACCTATGCGCACAGAGTTTGACCTTTGAGCGTGAAGAGTAT	67	
QY	754	gttctctatcccaaggttgaagttccataagaanaaagaatttgtcgagatggtacactt	81	
Db	678	GTCCCTTTTGCCAAAGGAGATGTGCACAAGAAAGAAATCATACAGATGTACCTTG	73	
QY	814	catgagccttgatgagcaagaatgcttcaagccacaagtgtagcaagatatltgtccctatg	87	
Db	738	CACGACTTGGAGCTGGCCAAATGCGGGCTGAGGCTGGGCAAGATATTCTGTCTATGTG	79	
QY	874	ggccagatgatatgaacccagaaagatctnaatcccgaaaaacacagcgccaaagatlaat	93	
Db	798	GGCCAGTTGTATGAACCCAAAAGAACAGAGATTCACAAATTAACCTTCAGAGGGGAGATCAAC	85	
QY	934	aagtggttaaalagatatatacgtatgaaagaaatltgcaagcttgtaccgtgtgttctc	99	
Db	858	AAGGTGTGAACCAATACATTGACACAGGCGCTTCAGAGCTGGTCCGTGACTCTT	91	
QY	994	attgatgaggtcccaatggttgatataagaatgttttcttatcttaacggttcattggag	105	
Db	918	GTTGACGAGGTCCACATGCTGGATGATGAGTGGCTTTTACCTACCTGACACCGAGCCCTGGAG	97	
QY	1054	agcccatatataccaatcgtgtatattcgtcacaatlaagggaaatgtatgtatgaagaag	111	
Db	978	TGCTTCATCGGCCCCCATTTGATCTTTGCATTCACACGAGAGCAACTGTGTGATCAGGGCC	103	
QY	1114	ac-----gatatagaagatccacatggtatatacggtagatcttctagatagctggtgatt	117	

[illegible]





QY 1054 agccattacaccatctgactgctacaaatagggagaaatgtaatgtaagaga 1113  
 DB 991 TCCCTCATGCCCCCATTTGTCATCTTGCATCAACCGAGCACTGTGCATCAGGGGC 1050  
 QY 1114 ac---tgatatagaacatccatggtatccgggtgatactctcctagatagctggtgatt 1170  
 DB 1051 ACCGAGGACATCTTCTCCACACGGGATCCCTTGACCTGCTGGACCGGGGTGATGATC 1110  
 QY 1171 atccgacagagacataatgacccctactgagatgatacatatgctactccagacaa 1230  
 DB 1111 ATGAGGACCATGCTGTATACGCCACAGAGATTAACGATCTTAAATTCGAGCCAG 1170  
 QY 1231 gtgagagatgatatgatatgataagaagtcgtcttatttagcgagatcgagacag 1290  
 DB 1171 ACGGAAGCATCAACATCACTGAGGAGGCCCTTAACCACTCGGGAGATGTGACCAAG 1230  
 QY 1291 acatcttgagacatgctatcattgatactccctggcggcggtgctcctaagaactat 1350  
 DB 1231 ACCACGCTGAGATTCAGTACGCTGACCCCTGCAACCTGCTGGCCAGATCCMAC 1290  
 QY 1351 ggaagagagaatctgcaagctgactcgaagaaagtcagtggtctatattgagatgcc 1410  
 DB 1291 GGGAGAGACACATTTGAGAGGAGCAGTGAGAGATCAGCGACTCTTCTATGACGCC 1350  
 QY 1411 aaatccctcgctcggtcgtccagaggaacaaagaagatacatcactgattgagatc 1470  
 DB 1351 AAGCTCTCCGCAAGATTTGCGCCGACCGACGAGCAAGTACATGATGAGTACGAGTT 1410  
 QY 1471 t 1471  
 DB 1411 t 1411

## RESULT 12

AF218072

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AF218072 1371 bp mRNA VRT 02-JAN-2001  
 Xenopus laevis pontin mRNA, complete cds.  
 AF218072.1 GI:12004635  
 African clawed frog.  
 Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 1371)  
 Etard, C., Wedlich, D. and Kuehl, M.  
 Xenopus laevis pontin, a Xenopus homolog of the prokaryotic RuvB  
 helicase  
 Unpublished  
 2 (bases 1 to 1371)  
 Etard, C., Wedlich, D. and Kuehl, M.  
 Direct Submission  
 Submitted (21-DEC-1999) Biochemistry, University of Ulm,  
 Albert-Einstein allee 11, Ulm D-89081, Germany  
 Location/Qualifiers  
 1..1371  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 1..1371  
 /note="X-pontin"  
 /codon\_start=1  
 /product="pontin"  
 /protein\_id="AA044127.1"  
 /translation="MKIEVSTTKTORIATHSHVGLGIDENGIAKOAAAGLYGOEN  
 AREAGVIVELKSKMAGRAVLLAGPVGKTKTALALIAELGKAPFCPWSEVY  
 SNEIKTEVLENFRAIGRLIRETEVEGEVELEIPCELENPKVGKTIASHVIG  
 LKTACTKOIKDPSIYSIOKREVEVDVITYEINSAGVNRQGSVDYATFEDLEA  
 ETVPLPKGVHOKREVIDYDITLHOLDVANARPOGGDILSMGOLMKPKTEIDKLR  
 GOINKVVKYIDOGIAELVPEYLFIDEVHMDIDRECFYTHALAESILAPIVATNRG  
 NCIRGTEDVASPHGIPLDLDKRWIITFMILTPEMKOITIKRAOTSGINISEALN  
 HLGIEIKTTLRYSVOLLTPANLAKINGKDSIEKHEVEINELFYDAKSSAKILAEQ

BASE COUNT 433 a 293 c 371 g 274 t  
 ORIGIN

Query Match 32.2%; Score 615.6; DB 8; Length 1371;  
 Best Local Similarity 66.2%; Pred. No. 7.3e-105;  
 Matches 904; Conservative 0; Mismatches 459; Indels 3; Gaps 1;

QY 94 atgagatcgagagagtgatcgatccctcgaagaagcagcgacccacacaccac 153  
 DB 1 ATGAAATCGAGAGAGTGAAGAGCACACAAAGCTACGCCATCGCTACCAAGCAT 60  
 QY 154 atcaagagacccgctcgaacccaatggaatggcagatgctgtgagcggtgctg 213  
 DB 61 GTGAAGGCGCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 214 ggcagcg 273  
 DB 121 GGCAG 180  
 QY 274 atggc 333  
 DB 181 ATGGCTGGAG 240  
 QY 334 ctgcgacatagcccgagagctcgcgcaaggtccctctctctctctctctctct 393  
 DB 241 TTAGCCATTGTCAGAGAACTGGGCAACAAAGTCCCTTTTGGCCATGGTGGCA 300  
 QY 394 gtgtactcctcgagagtgcaagaacctgaggtgctgagtgaaatcttcgtagag 453  
 DB 301 GTCTATTCCACAGATGATCAAGAAACCCAGAGTACTGATGAGAAATTTCCGAG 360  
 QY 454 ggttcgctataaagaagaagaagaaggttctgaagaaggttactgaacttccca 513  
 DB 361 GGACTGCGGATTAAGAGACTAAGAGAGTGTATGAAGAGAGAGTGAAGAGTCC 420  
 QY 514 gaagagctgagagtgatcaactgtgtgatatgcaaaaagattagcagatgata 573  
 DB 421 TGTGAGACAGAGATCCAAATGGAGAGATATGGCAACACATCGATCTTTATC 480  
 QY 574 ttaagagcttgaagagagactgaagcaactgaagttgacttctcaattatgag 633  
 DB 481 CTGAAACTCCAAAGAGAACCAACCAAGCTTAACCTTGAATCTATGAGATCTTA 540  
 QY 634 atcaagaaaggttgcgagtggtgagtgatgatatcatcgaagaacaatgtgag 693  
 DB 541 CAG 600  
 QY 694 aaaaagttgtagatgtgtctctctctctctctctctctctctctctctctct 753  
 DB 601 AAGCGCAAGGCGGATGTGATACATGACGCCAGCAATTTGACCTTGAGGTGAG 660  
 QY 754 gtctctatcccaaaagtgatgacatcacaagaagaagaagaatgtgagagatg 813  
 DB 661 GTCCCTCTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 814 catgaccttgatcagcaaatgtctcagccacagaggtggtgcaagaatgtgtcc 873  
 DB 721 CATGACCTGGATGTAGCCCAATGCGACGCGACAGGAGAGAGAGAGAGAGAG 780  
 QY 874 ggcagatgtagaagaacgaagaagctgaataccgcaaaactcagcgaagaat 933  
 DB 781 GGACAGCTCATGAAGGCCACAGAAACAAATTTACATTAAGCTGAAGAGGCAAT 840  
 QY 934 aaggtgtaatatagatalatcagatgaagaatgtgagagctgtgactgtgtgttc 993  
 DB 841 AAGAGGAGAAACAAATATCATGACCAAGGCACTGACACTTGTCCGGGGGTCT 900  
 QY 994 atgagtgatcacaagttgatatcgaatgtttcttctctctctctctctctct 1053  
 DB 901 ATAGATGAAGTACACATGTTGGATATAGAAATGTTTACATACCTCCACCGGCT 960







Wed Nov 14 08:34:08 2001

```

QY 1054 agccattatcaccaatcgtgatacttgcatacaaataggggaataatgtaagtgaagagga 1113
Db 1037 TCTTCTATCGCTCCCATCGTCATCTTTGTCATCCACCGAGGCAACTGTCATCAGAGGC 1096
QY 1114 act---gatagacaagtcacacatggtatcccggtggatcttctctagataggctggtgatt 1170
Db 1097 ACTGAGGACATCACATCCCTCAGGSCATCCCTCTTGACCTTCTGGACCGAGTGATGATA 1156
QY 1171 attcgagacagacatatggccctactgagatgatacacagatatggctatccgagacacaa 1230
Db 1157 ATCCGGACCATGCTGTATATCTCCACAGGAATGAACAGATCATTAATAATCCGTGCCAG 1216
QY 1231 gtgagagagattgatggatgagaaagtcttgccttatttagggagagatcggacagcag 1290
Db 1217 ACGGAGGAATCAACATCAGTGAGGAGGCACCTGAACCACTGGGGGAGATTGGCACCAAG 1276
QY 1291 acatctttgagacatgctattcaattgatataccctgccagcgtggtctcaagactaat 1350
Db 1277 ACCACACTGAGGTACTCAGTGCAGCTGCTGACCCCGCCAACTTGCTTAAATCAAC 1336
QY 1351 ggaagagagaaaatctcaagcgtgctctcaggaagtcagtggtgctctatttggatgcc 1410
Db 1337 GGAAGGACAGCATTTGAGAAAGACATGTCGAAGAGATCAGTGAACCTTTTCTATGATGCC 1396
QY 1411 aaatcctcggctcggctgctccaggaggaacaaagaaagatacatca 1456
Db 1397 AAGTCTCTCCGCCAAAATCCTGGCTGACCCAGCAGGATAAGTACATGA 1442

```

Search completed: November 13, 2001, 11:50:47  
Job time: 7047 sec